Stanford – Vienna
Transatlantic Technology Law Forum

A joint initiative of
Stanford Law School and the University of Vienna School of Law

European Union Law Working Papers

No. 115

The Regulation of Artificial Intelligence in the European Union Through GDPR and AI Act: Bias and Discrimination in AI-Based Decisions and Fundamental Rights

Sarah Buchheister

2025

European Union Law Working Papers

Editors: Siegfried Fina and Roland Vogl

About the European Union Law Working Papers

The European Union Law Working Paper Series presents research on the law and policy of the European Union. The objective of the European Union Law Working Paper Series is to share "works in progress". The authors of the papers are solely responsible for the content of their contributions and may use the citation standards of their home country. The working papers can be found at http://ttlf.stanford.edu.

The European Union Law Working Paper Series is a joint initiative of Stanford Law School and the University of Vienna School of Law's LLM Program in European and International Business Law.

If you should have any questions regarding the European Union Law Working Paper Series, please contact Professor Dr. Siegfried Fina, Jean Monnet Professor of European Union Law, or Dr. Roland Vogl, Executive Director of the Stanford Program in Law, Science and Technology, at:

Stanford-Vienna Transatlantic Technology Law Forum
http://ttlf.stanford.edu

Stanford Law School
Crown Quadrangle
559 Nathan Abbott Way
Stanford, CA 94305-8610

University of Vienna School of Law
Department of Business Law
Schottenbastei 10-16
1010 Vienna, Austria

About the Author

Sarah Buchheister studied law at Freie Universitaet Berlin and successfully completed her First State Examination in Germany. Recently, she obtained a Master of Laws (LL.M.) in European and International Business Law with distinction from the University of Vienna, Austria. During her master's degree, she also attended Stanford Law School as part of the Stanford Law, Science & Technology Program. She is currently starting a PhD in Berlin, while working as a research assistant at Freie Universitaet Berlin with a focus on European and Administrative Law.

General Note about the Content

Suggested Citation

This European Union Law Working Paper should be cited as:
Sarah Buchheister, The Regulation of Artificial Intelligence in the European Union Through GDPR and AI Act: Bias and Discrimination in AI-Based Decisions and Fundamental Rights, Stanford-Vienna European Union Law Working Paper No. 115, http://ttlf.stanford.edu.

Copyright

Abstract

This thesis examines the increasing importance of discrimination and bias in AI-based decision-making and decision support systems in their use by public authorities. Further, how the European Union's legal framework addresses this issue, with a focus on the GDPR and the AI Act. As AI technologies are increasingly used in public administration and due to their potential to influence outcomes in a discriminatory manner, this is a critical issue of our time. The research focuses on the associated risks to fundamental rights, which can be affected by biased algorithms and opaque decision-making processes, but also includes the major potential. Different areas of application in the Union are analyzed in more detail. The areas of law enforcement, EU border control, social welfare and the allocation of university places are discussed closely based on use cases. These use cases are subsumed under the legislation in focus, thereby identifying the scope of protection against discrimination by AI systems. While the provisions of the GDPR may provide certain fundamental protection against discrimination in advance, it does not actually offer comprehensive safeguarding. Art. 22 GDPR, which generally prohibits automated decision-making, is rather insufficient, especially due to its limited scope of application to solely automated processing. In part, the GDPR rather hinders effective protection against discrimination because of its strict requirements. Likewise, the AI Act does not consistently provide adequate protection of fundamental rights due to its risk-based regulatory system. While the legislator's aim of many provisions is to combat discrimination, its design as a product safety law prioritizes system-level security over the protection of individuals, a discrepancy that contrasts the individual-centric approach of the GDPR. Helpful provisions only apply to high-risk AI systems. Additionally, the risk category is self-assessed, thereby providing a limited level of protection. The study contributes to the understanding of AI-based decision-making and discrimination. It is in line with a number of critical positions regarding the inadequate protection provided by the legislation examined but goes beyond this with a nuanced case-based analysis.

Table of abbreviations

Abbreviation	Meaning
ACM Trans. Inf. Syst.	ACM Transactions on Information Systems
ADM	Automated Decision Making
AI	Artificial Intelligence
AMS	Arbeitsmarktservice
ANN	Artificial neural network
AöR	Archiv des öffentlichen Rechts
Berkley Tech. L. J.	Berkeley Technology Law Journal
Calif. Law Rev.	California Law Review
CFREU	Charter of Fundamental Rights of the European Union
CILA	Canadian Immigration Lawyers Association
Colo. Tech. L. J.	Colorado Technology Law Journal
Common Mark. Law Rev.	Common Market Law Review
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions
Comput. Law Rev. Int.	Computer Law Review International
Comput. L. & Secur. Rev.	Computer Law & Security Review
ed	editor
edn	edition
eds	editors
EJLT	European Journal on Law and Technology
ERA Forum	Journal of the Academy of European Law
Europol	European Union Agency for Law Enforcement Cooperation
Eur. Data Prot. Law Rev.	European Data Protection Law Review
Eur. J. Crim. Policy Res	European Journal on Criminal Policy and Research
Eur. J. Migr. Law.	European Journal of Migration and Law

Eur. Law J.	European Law Journal
EuZW	Europäische Zeitschrift für Wirtschaftsrecht
Fla. St. U. L. Rev.	Florida State University Law Review
Fordham L. Rev.	Fordham Law Review
FRA	European Union Agency for Fundamental Rights
FRONTEX	European Border and Coast Guard Agency
Ga. St. U. L. Rev.	Georgia State University Law Review
Ger. Law J.	German Law Journal
GPAI	General-Purpose AI Model
GPT	Generative pre-trained transformer
GRUR-Prax	Gewerblicher Rechtsschutz und Urheberrecht in der Praxis
Harv. Bus. Rev	Harvard Business Review
Harv. J. L. & Tech.	Harvard Journal of Law & Technology
iBorderCtrl	Intelligent Portable Border Control System
Int. Data Priv. Law	International Data Privacy Law
Int. J. Hum. Rights	International Journal of Human Rights
Int. J. Public Adm.	International Journal of Public Administration
Int. Rev. Intellect. Prop. Compet. Law	International Review of Intellectual Property and Competition Law
J. Bus. & Tech. Law	Journal of Business & Technology Law
JZ	Juristen Zeitung
MIT Tech. Rev.	Massachusetts Institute of Technology Review
ML	Machine Learning
MMR	MMR Zeitschrift für IT-Recht und Recht der Digitalisierung
Nat. Commun.	Nature Communications
NJW	Neue Juristische Wochenschrift

NVwZ	Neue Zeitschrift für Verwaltungsrecht
OECD	Organization for Economic Cooperation and Development
Precops	Pre-Crime Observation System
RDi	Recht Digital
Soc. Sci.	Social Sciences
Stan. L. Rev. Online	Stanford Law Review Online
Stanford HAI	Stanford University, Human-Centered Artificial Intelligence
SyRI	System Risk Indication
U. Pa. L. Rev.	University of Pennsylvania Law Review
UCLA L. Rev.	University of California Law Review
Univ. Ill. Law Rev.	University of Illinois Law Review
Va. J. L. & Tech.	Virginia Journal of Law and Technology
Va. L. Rev.	Virginia Law Review
Verfblog	Verfassungsblog
Wash. L. Rev.	Washington Law Review
Wm. & Mary L. Rev.	William & Mary Law Review
Yale J. L. & Technol.	Yale Journal of Law & Technology
Yale L. J.	Yale Law Journal
ZD	Zeitschrift für Datenschutz
ZEuS	Zeitschrift für Europäische Studien

Table of Contents

Introduction

"Artificial Intelligence is developing fast. It will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine."[1] This is how the European Commission aptly describes the development of Artificial Intelligence (AI). At the latest since ChatGPT was introduced at the end of 2022[2], AI has played a role in almost all areas of life and is becoming increasingly suitable for everyday use. In many areas, this technical progress leads to positive effects[3]. As such, AI also has an important function in automated decision-making, since tasks can be performed at a pace and precision that human labor cannot achieve. Yet, its implementation, especially when it comes to AI based decision-making harbors a certain risk with a view to fundamental rights, and particularly with regard to equality and non-discrimination.

Even before today's boom of generative AI, the use of algorithms and AI in several areas, both in the private and the public sector, led to questions about discrimination.[4] For a while, Amazon tested an algorithmic recruitment tool that turned out to be biased against women and therefore had to be discontinued.[5] Likewise, there was a stir about

[1] European Commission, 'White Paper on Artificial Intelligence - A European approach to excellence and trust' COM (2020) 65 final, p. 1.
[2] OpenAI 'Introducing ChatGPT', (*OpenAI*, 30 Nov 2022) <openai.com> accessed 2 Dec 2024.
[3] In general, e.g. 'Artificial intelligence: threats and opportunities' (*European Parliament* 20 June 2023) <europarl.europa.eu> accessed 2 Dec 2024.
[4] Use e.g. in the areas of labor market policy, social welfare, education, policing and justice, see Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 78 ff.
[5] See Jeffrey Dastin, 'Insight - Amazon scraps secret AI recruiting tool that showed bias against women' (*Reuters* 11 Oct 2018) <reuters.com> accessed 2 Dec 2024.

an algorithm relating to the Apple credit card, which granted men higher credit limits than women of equal financial standing.[6] It is not only private companies that implement algorithms, but also state agencies. For instance, in 2016 it was revealed that the algorithm COMPAS used in the United States to calculate, among other things, the risk of recidivism among prisoners, wrongly classified black defendants as more likely to reoffend than white defendants.[7] These examples illustrate that the use of algorithms and likewise the use of AI can lead to discriminatory and biased results. Therefore, the question arises to which extent the risks of discrimination and bias in AI-based decisions are already addressed, remedied and prevented under the current European Union legal framework.

To answer this question, a general understanding of the relevant terminology and technologies will be provided first. In order to gain a fundamental understanding, terms such as algorithms, rule-based- and machine learning, big data or AI-based decision-making are introduced and explained in the first chapter. This is followed by a classification of the central concepts of bias and discrimination. After this introductory part, areas of application for AI in the context of automated decision-making, both by the European Union and its Member States, will be examined. A focus will be placed on the application of AI in law enforcement, in border control and in the administrative sector. This is followed by an analysis of the reasons for the problems identified in relation to discrimination and bias in AI, to understand what legal approaches are needed. Based on this, a review of the current legal framework in the EU will be

[6] See Alisha Haridsani Gupta, 'Are Algorithms Sexist?' (*The New York Times* 15 Nov 2019) <nytimes.com> accessed 2 Dec 2024.
[7] See Julia Angwin, Jeff Larson, et al., 'Machine Bias' (*propublica* 23 May 2016) <propublica.org> accessed 2 Dec 2024.

provided, focusing on the General Data Protection Regulation[8] and the recently adopted AI Act[9] with a view to how these contribute to the prevention of discrimination. The most important provisions in both laws are considered, analyzed and interpreted in light of the use cases presented, to determine the extent to which the provisions can prevent discrimination in advance. On the grounds of the findings, both weaknesses in the legal framework and improvements are discussed. Going further, possibilities, both legal and technical, will be identified and presented to counteract this problem.

[8] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.
[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) [2024] OJ L 2024/1689.

I. Introduction to the basics of Artificial Intelligence

Eventhough not being a brand-new technology, AI has only recently become an integral part of our lives. Although we have probably all dealt with AI at some point, for example when using Siri, Alexa or even translation programs[10], the complexity of how it works is usually unknown. Therefore, the aim of this first chapter is to provide background information, as the law can only be properly developed and used if there is at least a basic understanding of the regulated topic. To build up this framework of understanding, the basic terms related to AI will be clarified and defined, along with a brief history of its development.

1. Algorithms and Artificial Intelligence

To begin with the central term algorithm. Algorithms are by no means a new technical development, but have been known for a long time.[11] An algorithm can be defined as a specific defined procedure that takes a value or a set of values as input and generates a value or a set of values as output in a specific period of time.[12] "An algorithm is thus a sequence of computational steps that transform the input into the output.".[13] In computer science, the term algorithm refers in general to a well-defined set of steps to achieve a specific goal.[14] Others compare an algorithm for simplified illustration with a cooking recipe, as it uses a rule-based step-by-step procedure to

[10] Siri and Alexa are AI-supported virtual assistants developed by Apple and Amazon; AI- powered translation programs are e.g. Google Translate or DeepL.
[11] The term originates from a translation of a writing by the Persian scholar Muhammed al-Khwarizmi, see Nele Heise, 'Algorithmen' in Jessica Heesen (ed), *Handbuch Medien-und Informationsethik* (J.B. Metzler 2016) p. 202.
[12] See Thomas H. Cormen, Charles E. Leiserson, et al., *Introduction to algorithms* (2nd edn The MIT Press 2022) p. 5.
[13] Ibid.
[14] Joshua A. Kroll, Joanna Huey, et al., 'Accountable Algorithms' (2017) 165 U. Pa. L. Rev. 633, 640.

solve a problem.[15] A simple example are algorithms that sort or order certain things. For instance, the input would be random numbers, the command or rule that the algorithm follows is sorting by size, thus the output would be a list of numbers sorted by size. Evidently, an algorithm can perform such a task significantly faster and even more accurately than a human being, making the use of algorithms a major advantage for such repetitive or organizing tasks.

1.1 Rule-based algorithm vs. Machine-learning algorithm

An important distinction, also with regard to AI, has to be made between rule-based algorithms and machine-learning algorithms. These two algorithms operate differently and therefore also have different areas of deployment.

1.1.1 Rule-based Algorithm

The main difference, which can already be recognized by the name, is that the rule-based algorithm works based on a defined set of rules, whereas a machine learning algorithm can self-learn and does not require to be programmed with predefined rules. With rule-based algorithms, the above comparison to recipes is quite suitable, as all rules are predefined, and all possible outputs are programmed in. This type can therefore be described as static, because the rules cannot change on their own, rather it always requires a human being to implement the new rules into the algorithm.[16] As a result, the traceability of the solution path of this type of algorithm is clear and can

[15] Uwe Kischel, 'Diskriminierung durch Algorithmen' in Volker Epping and Christian Hillgruber (eds), *BeckOK GG* (58th edn C.H. Beck 2024) para. 218a.
[16] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 33.

be understood.[17] Rule-based algorithms can therefore be used if the underlying logic is simple and clear and the decision-making process is rather straightforward.

1.1.2 Machine-learning Algorithm

Machine-learning algorithms are of greater relevance and at the same time more complex, because, as already mentioned, they can self-learn. In comparison to the rule-based algorithm, it can be described as dynamic, as the rules can change without an implementation by a human-being[18]. One of the well-known and basic examples of machine learning is the email spam filter.[19] Taking this spam filter example, a machine learning algorithm is trained to recognize patterns in emails flagged as spam, such as certain phrases or sender locations. Over time, the algorithm refines its ability to recognize spam with greater accuracy as it analyzes more data and may then automatically classify new emails based on the patterns it learned. This shows how machine learning improves its performance by continuously learning from additional examples.[20] This type of algorithm is directly related to AI, since they are the part of an AI which enables it to learn from data and experience and thus to improve its performance over time.[21] Therefore, to a certain extent, machine learning is the key element of an AI's success.[22] Accordingly, machine learning can be described as a sub-area of Artificial Intelligence.[23] In order to provide a clearer understanding, the following illustration is presented.

---

[17] Harry Surden, 'Artificial Intelligence and Law: and Overview' (2019) 35 Ga. St. U. L. Rev. 1306, 1316.
[18] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 33.
[19] Harry Surden, 'Artificial Intelligence and Law: and Overview' (2019) 35 Ga. St. U. L. Rev. 1306, 1312ff.
[20] Harry Surden, 'Machine Learning and Law' (2014) 89 Wash. L. Rev. 87, 90ff.
[21] Ibid, 89.
[22] Mario Martini, *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz* (Springer 2019) p. 20.
[23] Wolfgang Ertel, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021) p. 3.

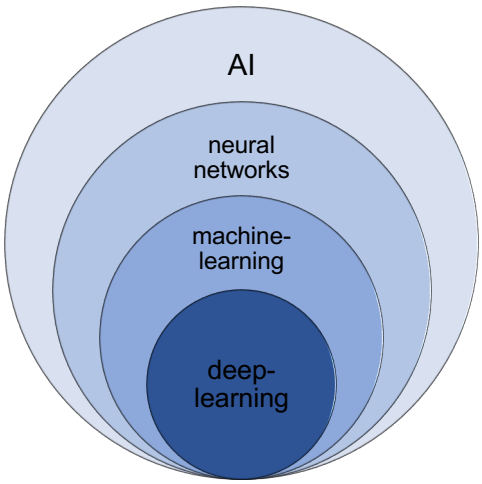

Illustration 1: Components of AI[24]

1.2 An attempt to define Artificial Intelligence

This raises the compelling question of what exactly AI is. Coming up with a precise definition for the increasingly used term AI is far more difficult than it is for the already known term algorithm. Firstly, it should be noted that AI is not a specific algorithm, but rather a field of computer science that uses - among other things – algorithms. AI therefore does not mean a specific tangible thing, instead it refers to a current technological development with the use of automated decision-making and machine learning.[25] It includes a wide range of concepts and terms, making it challenging to define.[26] AI can generally be described as the "capacity of computers or other machines to exhibit or simulate intelligent behaviour".[27] The definition, which is considered to be the first, originated from the 1995 paper 'A Proposal for the Dartmouth

[24] Illustration based on Ralf T. Kreutzer, *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmenseigene KI-Journey* (2nd edn Springer 2023) p. 10; see also John D. Kelleher, *Deep Learning* (MIT Press 2019) p. 6.
[25] FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022), p. 18.
[26] Andreas Häuselmann, 'Disciplines of AI: An Overview of Approaches and Techniques' in Bart Custers and Eduard Fosch-Villaronga (eds), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice* (T.M.C. Asser Press 2022) p. 44.
[27] 'Artificial Intelligence' (*Oxford English Dictionary* Dec 2023) <oed.com> accessed 2 Dec 2024.

Summer Research Project on Artificial Intelligence'.[28] There, John McCarthy defined the concept of AI as follows:

Artificial Intelligence is the science and engineering of making intelligent machines.[29]

A further definition that highlights the element of comparison with human intelligence is this one:

Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.[30]

Despite these and other numerous attempts to define AI conclusively, there is yet no generally valid definition. In a legal context, the question arises as to whether a specific definition of the term AI is necessary and expedient or whether, instead, more focus should be placed on the individual applications and the problems they explicitly contain.[31] Based on this approach, the EU has provided a fitting description in the recently enacted AI Act. Art. 3 of the AI Act defines 'AI systems' as follows:

'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

---

[28] John McCarthy, Marvin L. Minsky, et al., 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (1955) <stanford.edu> accessed 2 Dec 2024.
[29] Christopher Manning, 'Artificial Intelligence Definitions' *(Stanford HAI Sep 2020)* <hai.stanford.edu> accessed 2 Dec 2024; see also Wolfgang Ertel, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021) p. 1.
[30] Elaine Rich, *Artificial Intelligence* (McGraw-Hill 1983); Wolfgang Ertel, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021) p. 3.
[31] See e.g. Matthew U. Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies' (2016) 29 Harv. J.L. & Tech. 354, 359ff.

Such a broad definition does not preclude any AI applications and will also be suitable for future technologies that will certainly develop.[32] It also seems appropriate that in this definition no direct reference is made to the human intelligence, but rather the concepts of autonomy and inference are used. Even if the definition in the legislation could be considered too broad, regarding legal certainty[33] and too imprecise to differentiate between covered AI systems and other, non-covered systems[34] it still provides a framework for the classification as an AI.

Last but not least, here is ChatGPT's response to the prompt "Define AI in one sentence": '*AI (Artificial Intelligence) is the development of computer systems that can perform tasks typically requiring human intelligence, such as learning, problem-solving, and decision-making.*'[35] It becomes clear that even an AI itself does not offer a more precise definition than those already proposed. Nevertheless, a basic understanding and classification of the term is provided.

1.3 Elements and functioning of Artificial Intelligence (ML)

Relevant in this thesis will mostly be the machine-learning algorithms, which are in fact the core of the debate on legal regulation, as they are the relevant part of self-learning AI.[36] In the context of machine learning, there are three main methods for the training to take place. The methods are referred to as supervised, unsupervised and reinforcement learning.[37] Supervised learning is a type of machine learning in which

[32] See Recital 12 AI Act.
[33] See e.g. David M. Schneeberger, *Machine Learning in der Verwaltung: Rechtsfragen der Black-Box Problematik* (Verlag Österreich 2024) p. 56.
[34] Christiane Wendehorst, Bernhard Nessler, et al., 'Der Begriff des „KI-Systems" unter der neuen KI-VO' (2024) MMR 605, 610.
[35] OpenAI ChatGPT response to prompt 'Define AI in one sentence' (28 September 2024).
[36] See Phillip Hacker, 'Europäische und nationale Regulierung von Künstlicher Intelligenz' (2020) 30 NJW 2142, 2143.
[37] See Stuart J. Russell and Peter Norvig, *Artificial Intelligence: a modern approach* (4th edn NJ: Pearson 2021) p. 694ff.; another mixed form such as 'semi-supervised learning' is also often indicated, see p. 695.

an algorithm is trained on labeled data, i.e. the input data is paired with the correct output. The model learns to match the inputs to the desired outputs, enabling it to make predictions for new, unseen data.[38] In unsupervised learning, the algorithm receives unlabeled data and must find patterns or structures in it without explicit guidance. As a result, the model identifies hidden relationships or clusters in the data and assigns them accordingly.[39] Reinforcement learning is a different approach to machine learning in which a system learns to make decisions by interacting with its environment and receiving feedback in the form of rewards or punishments. Over time, it optimizes its actions to maximize long-term rewards and learns through trial and error.[40] In addition to these three main methods of machine learning, deep learning is of particular importance; this is not a separate method but can rather be used within them. Accordingly, the associated terms are briefly explained.

1.1.1 Deep Learning

Deep learning can itself be classified as a sub-area of machine learning, where the focus is on using deep neural network models.[41] Thus, "deep learning networks are neural networks that have many hidden layers of neurons".[42] These multiple hidden layers allow the network to learn complex patterns, making it suitable for tasks such as image recognition and natural language processing. With this type of machine learning, due to the multi-layered nature and the different levels of weighting of the connections, the decision-making process is not comprehensible and not transparent, resulting in

[38] Andreas Häuselmann, 'Disciplines of AI: An Overview of Approaches and Techniques' in Bart Custers and Eduard Fosch-Villaronga (eds), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice* (T.M.C. Asser Press 2022) p. 49f.
[39] Ibid, p. 50f.
[40] Stuart J. Russell and Peter Norvig, *Artificial Intelligence: a modern approach* (4th edn NJ: Pearson 2021) p. 695.
[41] John D. Kelleher, *Deep learning* (MIT Press 2019) p. 8.
[42] Ibid, p. 69.

the so-called 'black box' effect.[43] This in particular, creates one of the fundamental problems with automated decision-making, but more on this later.[44]

1.1.2 Neural Networks

Neural Networks are used for machine learning and as seen above, in the context of deep learning. They are computer models inspired by the structure and function of the human brain[45], although a far-reaching comparison between an artificial neural network and the human brain is subject to some criticism.[46] An artificial neural network imitates the structure of the brain as follows. It uses interconnected neurons to process data and consists of an input layer, hidden layers and an output layer. The ANN learns by adjusting the weights and biases in the neurons as it passes the data through these layers.[47] To make the difference to the human brain clear, the following analogy is very appropriate. Humans can distinguish between something that seems like an object and something that actually is that object - like a cloud that looks like a dog without actually being a dog. ANNs, however, do not have the ability to make such distinctions. Instead, they simply have to choose the label from their repertoire that best fits, so they would assume the cloud to be a dog.[48]

1.1.3 Big Data

Another key aspect is the data collection that is used to train an AI. This is because, in general, machine learning algorithms can only be as good as the data they are given

[43] Stuart J. Russell and Peter Norvig, *Artificial Intelligence: a modern approach* (4th edn NJ: Pearson 2021) p. 750.
[44] See III.2.2.
[45] John D. Kelleher, *Deep learning* (MIT Press 2019) p. 65.
[46] Only the basic characteristics remain comparable, the constant comparison can lead in part to placing human intelligence and artificial intelligence on the same level, see Aziz Z. Huq, 'A right to a human decision' (2020) 106 Va. L. Rev. 611, 637ff.
[47] John D. Kelleher, *Deep learning* (MIT Press 2019) p. 67.
[48] Zhenglong Zhou and Chaz Firestone, 'Humans can decipher adversarial images' (2019) 10, 1334 Nat. Commun. <nature.com> accessed 2 Dec 2024.

to analyze.[49] Due to this, the importance of data, and in particular of suitable training data, has increased enormously in recent years.[50] In connection with machine learning, and especially in the field of deep learning, the term big data is mentioned frequently. In general, big data refers to technological developments related to the collection, storage, analysis and application of data.[51] An exact definition of big data does not exist, instead it is often described by using certain characteristics: volume (size), velocity (speed of generation) and variety (different types of data like text, images, etc.).[52] Hence, these characteristics - the volume, the variety and the fast pace at which new data is collected - are the key factors that distinguish big data from 'normal' data.[53] Machine learning benefits from these large amounts of data because more data provides more examples from which the algorithms can learn. This again improves the accuracy and effectiveness of the models. Summed up, with big data, machine learning systems can recognize more complex patterns, make better predictions and gain more insights.

1.1.4 Data Mining

Within the field of machine learning and big data collections, there is often an association made with data mining. Data mining describes the automated process of extracting knowledge, information and correlations from large sets of data and further presenting that data by using algorithms.[54] This allows patterns or regularities in the

[49] Harry Surden, 'Machine Learning and Law' (2014) 89 Wash. L. Rev. 87, 106.
[50] See Mauritz Kop, 'The Right to Process Data for Machine Learning Purposes in the EU' (2021) 34 Harv. J. L. & Tech. 1,3ff.
[51] FRA, *#BigData: Discrimination in data-supported decision making* (Publications Office of the European Union 2022), p. 2.
[52] Giovanni Sartor, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 4; often enhanced by veracity (data quality) and variability (consistency of data over time), see e.g. FRA, *#BigData: Discrimination in data-supported decision making* (Publications Office of the European Union 2022), p. 2.
[53] Rok Dacar, 'The Essential Facilities Doctrine, Intellectual Property Rights, and Access to Big Data' (2023) 54 Int. Rev. Intellect. Prop. Compet. Law 1487, 1496.
[54] Wolfgang Ertel, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021) p. 206; Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges*

data sets to be uncovered, producing a series of calculated statistical relationships.[55] The main objective of this technique is to identify statistical relationships and thereby extract hidden patterns that are not immediately recognizable.[56] There is no clear cutoff between data mining and machine learning, yet the differences are rather fluid.[57] Data mining tends to focus on discovering hidden patterns and insights from large data sets, whereas machine learning develops models that can learn from data to make predictions or decisions, with an emphasis on generalization to new data.

1.4 The history of Artificial Intelligence - a brief overview

To conclude this first chapter regarding the technical basics, a brief outline of the history of AI is provided. The history of AI began in the 1950s, when pioneers such as Alan Turing and John McCarthy laid the foundations for it: Turing introduced the idea that machines could simulate human intelligence[58], and McCarthy shaped the term "artificial intelligence" as already seen above.[59] Early research regarding AI focused on symbolic thinking and problem solving. In the 1980s and 1990s, AI research made progress with expert systems and rule-based models, for instance, the first computers succeeded in winning against the best humans at chess.[60] The 2000s brought a revolution with machine learning and neural networks that enabled AI to learn from large data sets, driven by the rise of big data and faster computing power.[61] Today, AI

and opportunities for gender equality and non-discrimination law (Publications Office of the European Union 2021), p. 33.
[55] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 13.
[56] Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 677.
[57] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 14.
[58] Alan Turing, 'Computing Machinery and Intelligence' (1950) 49 Mind 433-460.
[59] John McCarthy, et al., 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (1955) <stanford.edu> accessed 2 Dec 2024.
[60] Wolfgang Ertel, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021) p. 8.
[61] Ibid.

has become widely used for applications such as self-driving cars, natural language processing and medical diagnostics, with deep learning playing the leading role.[62] Besides these areas of application in daily life, AI is also used for decision-making and decision support in the public sector to bring about fast and resource-saving decisions.

2. Generative Artificial Intelligence

Probably the latest development in the field of AI is the introduction of generative AI. Generative AI refers to all forms of AI that create new content such as texts, images, music or codes by learning patterns from existing data and thereby generating new, original results that did not exist before.[63] By modeling data distributions, it uses techniques such as Generative Adversarial Networks (GANs), transformers (as in large language models such as GPT) or Variational Autoencoders (VAEs) to generate new results that resemble the data on which it was trained.[64] The difference is therefore the creation of something new, whereas non-generative AI focuses on recognizing patterns, classifications and predictions without creating new content.

In the context of automated decision-making by AI, generative AI does not necessarily have a particularly key role, but it always depends on the area of application. Typical decisions that could be made by AI systems do not necessarily require the creative advantage that generative AI provides. For these reasons, the focus here and in the use of AI systems in decision-making is less on generative AI and more on the use of non-generative AI systems.

[62] For an overview see Ibid, p. 6ff.
[63] Sara D'Onofrio, 'Generative Künstliche Intelligenz – die neue Ära der kreativen Maschinen' (2024) 61 HMD Praxis der Wirtschaftsinformatik 331, 333.
[64] Ibid, 333ff.

II. Bias and Discrimination – A conceptual classification

The terms discrimination and bias are of central importance for this thesis. Therefore, they are explained and set in relation to each other and to the concept of fairness for a clearer understanding before examples of discriminatory AI technologies are given.

1. Bias

The term bias can initially be understood as an act of unfairly supporting or opposing a particular person or cause because one's judgment is influenced by personal opinions.[65] However, the fundamental concept of bias does in principle not imply anything negative; on the contrary, the described tendency can also be positive.[66] Within this context, though, it is precisely the unfair outcome as already given in the first definition that is important. It is a "differential treatment based on protected characteristics"[67], for instance ethnic origin, gender, religion, skin color or sexual orientation. This means that bias is a systematic preference or prejudice against a person, group or thing that can distort judgment and lead to unfair outcomes. The term is besides this also used in the context of statistical bias. Statistical bias is a systematic inaccuracy in data collection or analysis that leads to inaccurate or misleading results, often due to issues with data measurement.[68] Such statistical biases lead to consistently biased results in a particular direction and affect the overall validity of the conclusions that can be drawn from the data.[69] This type of bias can therefore have an effect on a result that derives from a data set.

[65] 'Bias' (Cambridge Dictionary) <dictionary.cambridge.org> accessed 2 Dec 2024.
[66] Batya Friedmann and Helen Nissenbaum, 'Bias in Computer Systems' (1996) 14 ACM Trans. Inf. Syst. (TOIS) 330, 332.
[67] FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022), p. 23.
[68] Ibid.
[69] Ibid.

In the context of this thesis, bias plays a role because it can also occur in algorithms. These algorithmic biases arise when algorithms make decisions in a way that systematically disadvantage certain groups of people.[70] Since the algorithms are based on historical data, they can maintain pre-existing human biases and reflect the biases in the data on which they were trained. Specifically, the often applied pattern recognition technology can unintentionally lead to the reproduction of human bias in various ways and its processing in the algorithm.[71] Accordingly, biases in algorithms are usually either due to human biases represented in the data set or to an over- or underrepresentation of certain groups in the data collection.[72] In part, bias in this regard is defined as computer systems that *systematically* and *unfairly* favor certain individuals or groups and disadvantage others.[73] For discrimination to be considered bias, it must therefore occur systematically and not just in isolated cases.[74]

2. Discrimination vs. Fairness

The question remains as to how these algorithmic biases are related to discrimination. The terms are closely linked, but they do not always occur side by side. When algorithmic bias disproportionately affects certain groups based on certain protected characteristics such as ethnicity, gender or age, it can lead to discrimination. However, not every bias leads to discrimination. It therefore depends on the characteristic by which an AI differentiates. For example, an algorithm that differentiates between people based on whether they own a pet does not directly refer to a protected

[70] Simon Friis and James Riley, 'AI And Machine Learning: Eliminating Algorithmic Bias Is Just the Beginning of Equitable AI' (2023) Harv. Bus. Rev. <hbr.org> accessed 2 Dec 2024.
[71] Robin Allen and Dee Masters, 'Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making' (2020) 20 ERA Forum 585, 588.
[72] Maya C. Jackson, 'Artificial Intelligence & Algorithmic Bias: The Issues with Technology Reflecting History & Humans' (2021) 16 J. Bus. & Tech. L. 299, 305ff. with additional references.
[73] Batya Friedmann and Helen Nissenbaum, 'Bias in Computer Systems' (1996) 14 ACM Trans. Inf. Syst. (TOIS) 330, 332ff.
[74] Ibid.

characteristic, and pet ownership is hardly a proxy for these protected characteristics.[75] As can be seen from this, not all bias in algorithms necessarily lead to discrimination. Rather the term bias is much broader than the term discrimination, as it encompasses any kind of disadvantage that can be considered wrong from an ethical or moral perspective regardless of a specific category.[76]

Within the legal framework of EU law, non-discrimination law takes on a significant role and is reflected in every form of European legislation. Art. 2 TEU[77], which lists the fundamental values on which the EU is based, states that equality is one of these.[78] The concept of discrimination is understood in the European legal context as follows. Discrimination is a disadvantageous and unjustified unequal treatment of persons or groups of persons in connection with a protected characteristic.[79] These protected characteristics to which algorithmic discrimination can refer, are for example those explicitly mentioned in Art. 19 TFEU[80]: sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation. Besides that, Art. 21 CFREU[81] prohibits any discrimination based on grounds 'such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation'.

In connection with algorithmic discrimination or discrimination by AI-supported systems, the terms fairness and further algorithmic fairness are used frequently and

[75] FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022), p. 24.
[76] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 47.
[77] Treaty on the European Union [2012] OJ C 326/13.
[78] The Unions objective of combating discrimination is also mentioned in Art. 3 (3) TFEU and Art. 10 TFEU.
[79] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 23 with additional reference.
[80] Treaty on the Functioning of the European Union [2012] OJ C 326/47.
[81] Charter of Fundamental Rights of the European Union [2012] OJ C 326/391.

mostly in the English-speaking area. There is no general understanding or even a definition of fairness, as everyone has differing approaches to what constitutes a fair decision, making it difficult to grasp the term in a legal context.[82] Yet, the concept of fairness can be distinguished in comparative and non-comparative fairness. In the comparative conception of fairness, fairness is assessed by comparing individuals or groups to ensure equal treatment and equal outcomes based on similar circumstances. In contrast, the non-comparative concept focuses on the intrinsic justice of actions or rules and assesses fairness regardless of how individuals are compared to each other, even if this leads to unequal outcomes.[83] With a view to the relevant European law, focusing on comparative fairness is more appropriate, as it works in a similar way to EU non-discrimination law, where comparative categories or groups are also formed. This means that the comparison is carried out between people or groups that have both been scored or assessed by the same AI-system. Algorithmic fairness presupposes that the score determined by a particular algorithm should be equally accurate for people belonging to a comparable group. This can be assumed if the ratio between the false-positive rate and the false-negative rate is the same for the relevant groups evaluated by an algorithm.[84] It becomes clear that fairness in principle means a condition in which there is no discrimination. In other words, fairness encompasses the legal concepts of equality and non-discrimination, but also goes beyond. The principle of equality in EU law appears to be more narrowly defined than the broader concept of fairness.[85] In order to create a concrete point of reference in the following,

[82] Thomas B. Nachbar, 'Algorithmic Fairness, Algorithmic Discrimination' (2021) 48 Fla. St. U. L. Rev. 509, 523.
[83] See Deborah Hellman, 'Measuring Algorithmic Fairness' (2020) 106 Va. L. Rev. 811, 834f.
[84] Ibid, 835ff.
[85] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 48.

and since the concept of fairness is unfamiliar to EU law, the following analysis is based on algorithmic AI discrimination.

III. Artificial Intelligence based decision-making

As already seen, the use of AI is broad and is implemented in numerous everyday situations. For instance, AI can create texts and images, monitor the home or recommend the next travel destination, however, the present thesis - with regard to the risks of discrimination to be examined - focusses on AI in the use of decision-making.[86]

AI-based decision-making relates to the use of AI systems to make or support decisions by analyzing data, identifying patterns and providing insights or actions. This approach makes it possible to automate or support complex decision-making processes that traditionally require human judgement. The advantages of using AI systems in decision-making are obvious: savings in labor resources and time, as well as supposedly objective and consistent decision-making. However, there are not only advantages to its use, but it also carries risks. There can be a significant risk to fundamental rights, and especially to equality rights, associated with the use of AI in forecasting and decision-making. When AI supports or even replaces the human decision, it can lead to discriminatory results as seen in the introductory examples. Although algorithms are often seen as objective tools for decision-making, machines are not inherently neutral as they are created and operated by humans, meaning biases present in human decision-making can be passed on to the algorithms and the systems created by them.[87] However, this begs the question of whether this does not

[86] For an overview, see Ibid, p. 78 ff; see also Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 30ff.
[87] FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022), p. 17.

also offer the possibility of enabling decision-making entirely without bias and discrimination, if only the underlying algorithms were properly trained? Further discussion on this and other approaches to addressing the problem will follow later.

With a look to ADM a distinction is to be made between fully automated decision-making and partly automated decision-making. Fully automated decision-making refers to scenarios where decisions are made entirely by automated systems without human intervention, such as an AI algorithm approving loan applications based solely on predefined criteria.[88] In contrast, partly automated decision-making involves a combination of human input and automated processes where the final decision requires human judgment or oversight, such as an algorithmic system used to assess an individual's creditworthiness employing AI while requiring a human to make the final decision.[89] Although it is a legally relevant distinction, as it will be seen below with regard to Art. 22 GDPR, apart from this it may be rather irrelevant. This is because even if it is only a partly automated decision, the direction or suggestion given by the algorithm is usually followed without the human taking their own additional criteria into account.[90] For this reason, no distinction between the two types is made in the following part of the examples.

In order to comprehensively illustrate the challenges of AI based decision-making in the context of discrimination, use cases in areas that are particularly relevant to fundamental rights and therefore examples in the public sector will be presented.

[88] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 21ff.
[89] See Frederik J. Zuiderveen Borgesius, 'Strengthening legal protection against discrimination by algorithms and artificial intelligence' (2020) 24 Int. J. Hum. Rights 1572, 1573.
[90] Frederik Zuiderveen Borgesius, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making* (Council of Europe 2018) p. 11.

1. Application of Artificial Intelligence in decision-making

Application of AI to support or replace human decision-making is emerging in numerous areas across the public sector.

1.1 Application in Law Enforcement

In the field of law enforcement, AI systems are used in various areas. The first example illustrating the challenges within ADM is the already mentioned COMPAS system, which is probably the best-known example regarding ADM and discrimination. Even though this does not take place in the EU yet, it needs to be presented as such a deployment in the EU is very conceivable in the future and highly relevant to fundamental rights. The COMPAS system is a risk assessment tool used in the United States criminal justice system to predict a defendant's likelihood of recidivism.[91] The defendant's risk of recidivism is assessed on the basis of factors such as previous convictions, socio-economic background and answers to a questionnaire to generate a recidivism-risk score between 1 and 10.[92] An investigation by ProPublica found that COMPAS predicted a disproportionately higher risk of recidivism for black defendants than for white defendants, even when factors such as prior criminal history weighed against this.[93] Black defendants were more likely to be categorized as high risk, while white defendants were more likely to be categorized as low risk. This racial bias has raised genuine concerns about discrimination in the application of the system. The company also responded to this criticism and explained that the same risk score leads

[91] Frederik Zuiderveen Borgesius, 'Strengthening legal protection against discrimination by algorithms and artificial intelligence' (2020) 24 Int. J. Hum. Rights 1572, 1574.
[92] Andrew Lee Park, 'Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing' (2019) UCLA L. Rev. <uclalawreview.org> accessed 2 Dec 2024.
[93] See Julia Angwin, Jeff Larson, et al., 'Machine Bias' (*propublica* 23 May 2016) <propublica.org> accessed 2 Dec 2024; Jeff Larson, Surya Mattu, et al., 'How We Analyzed the COMPAS Recidivism Algorithm' (*propublica* 23 May 2016) <propublica.org> accessed 2 Dec 2024.

to the same probability of recidivism, regardless of whether the person is black or white.[94] Overall, it led to an intense debate about fairness and the basis of judges' decision-making.[95] Apart from the discrimination, this example illustrates perfectly, how ADM is and can be used to support human decision-making in criminal justice.

AI-based decision-making further comes into play in the area of so-called predictive policing, which is already used in the Union.[96] Predictive policing is the application of analytical techniques to identify likely targets for police action and to prevent crime or solve past crimes through statistical predictions.[97] Predictive policing systems can also perpetuate and reinforce existing biases contained in historical crime data. If certain locations (predictive mapping) or certain people (predictive identification) have been over policed or treated unfairly in the past, the algorithms may categorize them as higher risk, leading to further discrimination and excessive policing.[98] This means if the police patrols non-white neighborhoods more frequently, crime statistics will show a disproportionate number of non-white offenders, as they are overrepresented in the monitored population. As a result, AI-supported predictive policing based on this data assigns a higher crime risk to non-white neighborhoods.[99] One example of an application in the EU is the Precobs system developed in Germany. This is a predictive policing tool that helps law enforcement agencies anticipate property crimes such as

[94] William Dieterich, Christina Mendoza, et al., 'COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity' (*Northpointe* 8 July 2016) <go.volarisgroup.com> accessed 2 Dec 2024.
[95] See Anne L. Washington, 'How to argue with an algorithm: Lessons from the COMPAS-ProPublica debate' (2018) 17 Colo. Tech. L. J. 131, 148ff.
[96] See FRA, *Getting the Future Right: Artificial Intelligence and Fundamental Rights* (Publications Office of the European Union 2020) p. 34ff.
[97] Walter L. Perry, Brian McInnis, et al., *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (Rand 2013) p. 1.
[98] Will Douglas Heaven, 'Predictive policing algorithms are racist. They need to be dismantled.' (2020) MIT Tech. Rev. <technologyreview.com> accessed 2 Dec 2024.
[99] Andrew Guthrie Ferguson, 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U. Pa. L. Rev. 327, 401ff.

burglaries by using historical crime data.[100] Another example is the Crime Anticipation System (CAS)15, that was developed in the Netherlands to combat crimes such as burglaries, robberies, and bicycle theft. It conducts weekly analyses using local, up-to-date data, enhanced with external information about neighborhoods and residents. Additionally, the system incorporates police findings on criminal activity, local conditions, and data from Dutch statistics.[101]

1.2 Application at EU borders

Another field where AI is more and more implemented is the European Union's border control.[102] Other countries have already been using AI systems to monitor borders for some time; Canada, for example, has been using automated decision-making systems in its immigration and refugee system since 2014.[103] The EU itself explains this approach with its "efforts to strengthen border control and mitigate security risks related to cross-border terrorism and serious crime".[104] Further, the deployment of AI at the EU's borders is intended to lead to a "smartening" of borders.[105] Systems that are already being used or tested in the EU include automated biometric systems, AI for emotion detection and algorithmic profiling.[106] One example is the so-called iBorderCtrl, an AI system that was tested at EU borders in Hungary, Greece and Latvia

[100] Ishmael Mugari and Emeka E. Obioha, 'Predictive Policing and Crime Control in The United States of America and Europe: Trends in a Decade of Research and the Future of Predictive Policing' (2021) 10 Soc. Sci., p. 7 with additional reference <mdpi.com> accessed 2 Dec 2024.
[101] Europol, *AI and Policing: The benefits and challenges of artificial intelligence for law enforcement* (Publications Office of the European Union 2024) p. 16.
[102] For an overview see Erik Silfversten, Luke Huxtable, et al., 'Artificial Intelligence – based capabilities for the European Border and Coast Guard' (FRONTEX 17 May 2021) <frontex.europa.eu> accessed 2 Dec 2024.
[103] See Gideon Christian, 'AI Facial Recognition Technology in the Canadian Immigration System' (*CILA* 29 August 2023) <cila.co> accessed 2 Dec 2024; Petra Molnar and Lex Gill, 'Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System' (*IHRP* 2018) <ihrp.law.utoronto.ca> accessed 2 Dec 2024.
[104] Costica Dumbrava, *Artificial Intelligence at EU borders: Overview of applications and key issues* (European Parliament 2021) p. 1.
[105] Ibid.
[106] Costica Dumbrava, *Artificial Intelligence at EU borders: Overview of applications and key issues* (European Parliament 2021) executive summary, p. 11ff.

to screen non-EU citizens through automated questioning and biometric verification. It uses AI-powered deception detection to analyze facial expressions and assess the likelihood of travelers providing false information, which raises concerns about accuracy and data protection.[107] Particularly in an area that is so sensitive to fundamental rights, the use of AI can quickly lead to results that have serious consequences for those affected. There are two main risks in the use of AI systems in border control. For sure, there is a high risk of discrimination when using AI in connection of people seeking protection at borders.[108] Furthermore, there is a limited possibility of challenging the biased automated decision due to the opaque decision-making process of the AI algorithm.[109]

1.3 Application in the Administrative Sector

AI systems offer great potential in the administrative sector to speed up, improve and simplify the process for both citizens and public authorities. There are numerous areas in public administration where the use of AI technology would bring benefits.[110] Here is an overview illustrating potential problems with the use of AI in public administration.

In Austria, an AI classification system has been used in Job Centers in the recent years which has attracted a great deal of attention and controversy.[111] The so-called AMS system refers to the Labour Market Service Austria (AMS). This system plays a key role in managing the labor market by providing a variety of services such as job

[107] See Fabio Chiusi, Sarah Fischer, et al. (eds), 'Automating Society' (*Algorithm Watch, Bertelsmann Stiftung* Oct 2020) p. 27ff. <algorithmwatch.org> accessed 2 Dec 2024.
[108] See Niovi Vavoula, 'Artificial Intelligence (AI) at Schengen Borders: Automated Processing, Algorithmic Profiling and Facial Recognition in the Era of Techno-Solutionism' (2021) 23 Eur. J. Migr. Law 457, 470ff.
[109] Mirko Forti, 'Addressing Algorithmic Errors in Data-Driven Border Control Procedures' (2024) 25 Ger. Law J. 635, 641.
[110] See Bernd W. Wirtz, Jan C. Weyerer, et al., 'Artificial Intelligence and the Public Sector - Applications and Challenges' (2019) 42 Int. J. Public Adm. 596, 600ff.
[111] See Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 43.

placement, career counseling, training programs and unemployment benefits. In recent years, AMS has integrated an AI-supported system that helps in the assessment of job seekers and the administration of labor market policy measures. The algorithm used is a scoring system that uses the already existing AMS database to predict the integration chances into the labor market, i.e. the employability of job seekers.[112] The system uses jobseeker data, as age, gender, nationality, education or previous employment, and assigns jobseekers to one of three categories on this basis (high/medium/low employability).[113] One of the main criticisms of the AMS algorithm is that it can reinforce existing social inequalities. Factors such as gender, age or nationality could unfairly disadvantage certain groups (e.g. older workers, women or immigrants), which leads to distorted results. For example, women are rated more negatively in general than men, furthermore points are deducted for women who may have caring responsibilities, whereas this circumstance is not viewed negatively for men.[114] However, Austria's AMS system, particularly the algorithm introduced to assess jobseekers, highlights both the potential and risks of using AI in public administration. While the system aims to increase efficiency and personalize support, it had raised concerns about discrimination, bias and transparency in its use.

Systems are further used around social welfare and within this context for the detection of fraud.[115] One system that has been identified to operate in a discriminatory manner is the SyRI system, a Dutch government initiative to detect fraud in welfare, tax and social security programs. SyRI used an automated algorithm to match and analyze

[112] Ben Wagner, Paola Lopez, et al., 'Der AMS-Algorithmus: Transparenz, Verantwortung und Diskriminierung im Kontext von digitalem staatlichem Handeln' (2020) 2 juridikum 191, 191.
[113] Ibid.
[114] Wiebke Fröhlich and Indra Spiecker, 'Können Algorithmen diskriminieren?' (*Verfblog* 26 Dec 2018) <verfassungsblog.de> accessed 2 Dec 2024; Barbara Wimmer, 'Der AMS-Algorithmus ist ein „Paradebeispiel für Diskriminierung"' (*futurezone* 17 Oct 2018) <futurezone.at> accessed 2 Dec 2024.
[115] See FRA, *Getting the Future Right: Artificial Intelligence and Fundamental Rights* (Publications Office of the European Union 2020) p. 30ff.

data from various government sources to identify individuals or households at risk of committing benefit or tax fraud.[116] It has mainly been used to detect fraud in areas such as welfare, health benefits, child benefit and tax evasion. The system was heavily criticized and eventually discontinued due to the concerns about discrimination and lack of transparency. The criticism was based on the system disproportionately targeting low-income neighborhoods where immigrants and marginalized groups lived. This led to concerns about discriminatory profiling, as the algorithm potentially reinforced biases based on socioeconomic status, ethnicity or geographic location. Many argued that SyRI unfairly focused on vulnerable groups and treated them as more at risk of fraud due to their socio-economic position.[117] Another issue was a lack of transparency, as the algorithm used in SyRI was a 'black box' - its decision-making process was not transparent. Citizens and even regulatory authorities barely knew how the system determined risk profiles or why certain people were flagged. As a result, it was impossible for people to find out whether they had been unfairly targeted or even to challenge the algorithm's decisions.[118] Due to these issues, the District Court of The Hague ruled that the system may no longer be used.[119] The court found that the government had not sufficiently justified the necessity of SyRI's data processing methods and that the system lacked adequate transparency and safeguards against discrimination.[120]

The final example is the application in the education sector and especially in areas such as university admissions and the allocation of study places. In France, a system

[116] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 81.
[117] Fabio Chiusi, Sarah Fischer, et al. (eds), 'Automating Society' (*Algorithm Watch, Bertelsmann Stiftung* Oct 2020) p. 162 <algorithmwatch.org> accessed 2 Dec 2024.
[118] Ibid.
[119] Rb. Den Haag, 5 Feb 2020, C/09/550982/HA ZA 18/388.
[120] Ibid, para 6.86f.

('Parcoursup') is used to allocate places at higher education institutions.[121] It uses an algorithm for the allocation of study places by assigning students to selected study programs based on a combination of student preferences, availability of study places and the assessment criteria of the respective university. The Parcoursup algorithm has been criticized because it uses data on applicants' income and place of residence in the allocation decision, which can lead to a discrimination against students from disadvantaged or rural areas.[122] Thus, the algorithm tends to favor students from well-known elite schools, especially from wealthier urban areas when evaluating applications. This means students from less prestigious high schools, particularly in rural areas or low-income neighborhoods, have a harder time competing.

2. The potential risks of Artificial Intelligence based decision-making

Based on the use cases presented two main risks stand out. Firstly, the risk of discrimination and secondly the lack of transparency.

2.1 Risks of discrimination

Through the given examples, it becomes clear that discrimination through AI based decision-making is a current and future-relevant problem that is anything but theoretical, even though the above examples are by no means all algorithm-based AI applications in which a bias has already been identified and lead to discrimination. The discrimination will mostly occur against certain groups, especially those who have already been discriminated in the past by humans, such as women, black people or migrants. After all, discrimination in AI based decision-making is and will become an

[121] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 44.
[122] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 61.

even greater significant issue as biased algorithms can disproportionately impact marginalized groups, leading to unfair treatment in critical areas such as employment, border decisions or criminal justice. In addition to these risks of discrimination for groups that have already suffered discrimination in the past, there is more to consider.

Even seemingly neutral data points can unintentionally lead to discriminatory results in algorithm-based decision making, due to the use of variables in algorithms that serve as indirect indicators of protected characteristics thereby serving as reliable 'proxies'.[123] This means that a generally neutral criterion used in the algorithm for the final result can also refer to protected grounds by proxy, for example class membership.[124] With data such as zip codes or even purchase history, it is conceivable that these correlate strongly with sensitive characteristics, leading to distorted results even if the sensitive characteristics are not directly included in the algorithm. An issue within this context, which is being discussed to some extent, is the emergence of new forms of discrimination in AI-based decisions.[125] Such new types of discrimination can affect individuals based on a wider range of criteria, including geographical location or even digital behavior. Thus, individuals who do not belong to traditionally disadvantaged groups may still be disadvantaged by algorithms because they recognize patterns based on different criteria.[126] A comprehensive look shows that both, historically already affected groups as well as new, previously non-discriminated groups can be subject to the risks if the algorithm-based AI recognizes correlations on which basis it then categorizes.

[123] Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 691ff.
[124] Ibid, 691.
[125] See Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising' (2020) 35 Berkley Tech. L. J. 367, 413ff.
[126] Ibid, 415.

2.2 Lack of transparency and the Black-box effect

An issue that occurs side by side with the risk of discrimination is the lack of transparency in AI based decision-making. The decision-making process of an AI, especially an AI that works on ML using neural networks, is not comprehensible, or to put it another way – "it is not explainable in human language".[127]

An algorithm may be mathematically comprehensible, but that is of no use to laymen, users or those affected by an AI-supported decision. Although experts understand the decision-making process to a certain extent if they can see the code and the inputs and outputs, they also reach their limits.[128] This resembles the so-called 'Black Box'[129] where the process and features which are involved in the decision-making are unknown or protected.[130] This Black Box effect plays a particularly significant role when using the deep learning method. The systems are programming themselves in a way that is incomprehensible to humans, even the person who built the system cannot understand the process.[131] This means that "deep learning is a particularly dark black box" by nature.[132] While the issue of these opaque algorithms in AI systems is a separate and independent problem, this thesis focuses on the risks of discrimination for which reason the black box problem is only considered in conjunction with it. As

[127] Tal Z. Zarsky, 'Transparent Predictions' (2013) 4 Univ. Ill. Law Rev. 1503, 1519.
[128] Mario Martini, *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz* (Springer 2019) p. 28; see also Joshua A. Kroll, Joanna Huey, et al., 'Accountable Algorithms' (2017) 165 U. Pa. L. Rev. 633, 657ff.
[129] The term was already used in 1969, see John McCarthy and Patrick J. Hayes, 'Some philosophical problems from the standpoint of artificial intelligence' (*Stanford Computer Science* 1969) <stanford.edu> accessed 2 Dec 2024; and is now commonly used for this problem, see e.g. Mario Martini, *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz* (Springer 2019) p. 29; see also Sandra Wachter, Brent Mittelstadt, et al., 'Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR' (2018) 31 Harv. J. L. & Tech. 841, 842ff.
[130] Joana Gonçalves-Sá and Flávio Pinheiro, 'Societal Implications of Recommendation Systems: A Technical Perspective' in Henrique Sousa Antunes, Pedro Miguel Freitas, et al. (eds), *Multidisciplinary Perspectives on Artificial Intelligence and the Law* (Springer 2024) p. 58.
[131] Will Knight, 'The Dark Secret at the Heart of AI' (*MIT Technology Review* 11 April 2017) <technologyreview.com> accessed 2 Dec 2024.
[132] Ibid, describes in depth how this is predisposed by the countless artificial neurons and intricately interconnected layers in a neural network for deep learning.

such, the problem with transparency around biased algorithms in AI based decision-making is firstly the difficulty in understanding how this bias is introduced and spread into the system and secondly how exactly the AI reached its decision, i.e. based on what criteria. Without clear insights into the data sources, model selection and decision-making processes, deployers cannot adequately assess or address the potential for discriminatory outcomes. This lack of transparency undermines public trust in AI systems, particularly in sensitive applications such as hiring or law enforcement, and potentially the most important point of legal protection for those affected by AI-based decisions.

Nonetheless, there are also opinions arguing AI-based decision-making is fundamentally no more opaque and no greater risk than human decision-making, i.e. there is no transparency gap.[133] Especially as the human decision-making and the basis for the decisions are not evident and so there exists likewise an opacity of other minds.[134] This approach is plausible and quite reasonable, however, some aspects speak in favor of a distinction that automated decision-making pose a greater risk than biased human decisions. Firstly, AI's efficiency and scope amplifies the impact of biases, as a machine processing numerous applications may apply biased criteria broadly to a lot more cases.[135] Secondly, AI systems can self-replicate biases, by reinforcing flawed outputs through self-feeding, which leads to entrenched stereotypes.[136] Lastly, human errors can be corrected through social and legal mechanisms, allowing for appeals and audits, whereas the opacity of AI technology makes it difficult to apply similar corrective measures to automated systems, as

[133] Aziz Z. Huq, 'A right to a human decision' (2020) 106 Va. L. Rev. 611, 611.
[134] Ibid, 643ff.
[135] Gabriele Spina Alì and Ronald Yu, 'Artificial Intelligence between Transparency and Secrecy: From the EC Whitepaper to the AIA and Beyond' (2021) 12 EJLT p. 5.
[136] Ibid.

providers often resist public scrutiny.[137] Accordingly, the problem of an discriminatory AI-based automated decision can certainly be seen as a greater risk compared to a human decision.

As mentioned above, the black box problem exacerbates the issue of discrimination, as it makes it even more difficult to ensure compliance with the existing law, which is intended to protect fundamental rights.

IV. Reasons for Bias and Discrimination

The reasons for this problem are various, but they mainly originate from the data and the structure of the algorithm. Additionally, opaque decision-making processes can obscure the sources, making it difficult to identify and rectify discriminatory outcomes.

1. Biased Data

Biases can stem from biased training data that reflects existing societal inequalities, leading algorithms to replicate and reinforce these prejudices.

1.1 Historical Bias

First, probably the most obvious and least technical cause is historical bias.[138] Historical bias is related to the presence of discriminatory practices or biases in historical data sets, that are used as training data. Bias and discrimination are deeply rooted in society and shaped by psychological, social and cultural factors. Already

[137] Ibid; see also Ivana Bartoletti and Raphaële Xenidis, *Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination* (Council of Europe 2023), p. 28ff.
[138] The term is used often to refer to this issue, see e.g. Aziz Z. Huq, 'A right to a human decision' (2020) 106 Va. L. Rev. 611, 686; Phillip Hacker, 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law' (2018) 55 Common Mark. Law Rev. 1143, 1148.

existing biases are maintained by the self-learning algorithms.[139] For example, if a hiring algorithm used in the labour market sector is trained based on historical employment data from a company that has systematically favored certain demographic groups, it may continue to favor those groups while disfavoring others. Existing biases are perpetuated and discriminatory outcomes in decision-making processes are reinforced as the algorithm learns from it and repeats the biased patterns in the historical data.[140] Notably, the algorithm effectively learns to make decisions that perpetuate historical inequalities. In effect, automating the process will turn any conscious and even unconscious biases or implicit biases of those involved in previous human decisions into a formalized rule which systematically changes the prospects of all individuals affected.[141] In other words, it can be described as follows - bias in, bias out - since the model only reproduces what is already present in the data.[142]

1.2 Under- and overrepresentation in Data

Another reason is the lack of data available for certain groups of the population, as underrepresentation in the data sets used as a basis for the training will also lead to discriminatory outcomes.[143] As a result of the missing data, also a high quality data set (i.e. the included data does not show any biases) will suffer from statistical biases when different groups are not represented in correct proportions.[144] For example, it has

[139] Phillip Hacker, 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law' (2018) 55 Common Mark. Law Rev. 1143, 1148.
[140] Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 681ff.
[141] Ibid, 682.
[142] See Sandra G. Mayson, 'Bias in, Bias out' (2019) 128 Yale L. J. 2218ff.
[143] See on this e.g. Maja C. Jackson, 'Artificial Intelligence & Algorithmic Bias: The Issues with Technology Reflecting History & Humans' (2021) 16 J. Bus. & Tech. Law 299, 305ff; see also Susanne Rentsch, ' „Computer sagt nein"- Gesellschaftliche Teilhabe und strukturelle Diskriminierung im Zeitalter Künstlicher Intelligenz' in Andreas Wagener and Carsten Stark (eds), *Die Digitalisierung des Politischen: Theoretische und praktische Herausforderungen für die Demokratie* (Springer 2023) p. 31ff. with additional reference.
[144] Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 684.

already become clear that women continue to be underrepresented in the data basis in almost all areas of life, such as healthcare, workplace policies or urban planning.[145] This gender data gap now continues to lead to systematic disadvantages in numerous areas for women, when AI systems are learning on its basis. Not only women but also other groups of people are suffering from this data problem. It affects all people who live on the "Big Data's margins"[146], e.g. due to poverty or geographical location, which means their lives may be less "datafied" than others.[147] Likewise, overrepresentation can also lead to discrimination, depending on the context in which the AI system is used. This is particularly relevant in the monitoring or screening of certain groups of people, which naturally leads to more infringements being detected among these people, however, only because of a more intensive investigation.[148]

1.3 Feedback Loops

'Feedback loops' are another issue in connection with biased data.[149] Feedback loops refer to settings in which the output of an AI system is fed back into the system as input, creating a cycle that can reinforce existing patterns or behaviors, i.e. the predictions made by the AI system influence the data used to update the same system.[150] This phenomenon can intensify the entire problem, as biased decisions lead back to data that reinforces this bias and perpetuates discriminatory outcomes.

---

[145] See Caroline Criado Perez, *Invisible Women: Exposing data bias in a world designed for men* (Vermilion 2020).
[146] Jonas Lerman, 'Big Data and its Exclusions' (2013) 66 Stan. L. Rev. Online 55, 57 <stanfordlawrevire.org> accessed 2 Dec 2024.
[147] Ibid.
[148] See Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 687.
[149] The Term 'feedback loops' is commonly used to describe this phenomenon, see e.g. Pauline T. Kim, 'Data-Driven Discrimination at Work' (2017) 58 Wm. & Mary L. Rev. 857, 882; also used in the AI Act, see Rec. 67 and Art. 15 (4) AI Act.
[150] FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022), p. 8, 29ff.

1.4 Proxies

As already described[151] in the context of the general risk of discrimination, unbiased data sets can still lead to discriminatory outcomes. This is due to the correlation between proxy characteristics the algorithm may link with other protected characteristics.[152] One good example of this are zip codes.[153] A zip code can serve as a proxy for a protected characteristic in an algorithm, e.g. race, ethnicity, socioeconomic status or gender, if geographic areas are strongly correlated with these characteristics. This can unintentionally lead to algorithmic discrimination, even if the algorithm does not explicitly take these protected characteristics into account. By this, it is becoming clear that removing sensitive and protected characteristics from the data set and thus from the basis for the decision is not enough to counteract the issue, as an AI system is able to find other related criteria. The main risk here is that this form of discrimination is subtle and difficult to recognize, as it often appears in seemingly neutral characteristics that unintentionally reflect existing social inequalities.

2. Structure of the algorithm

A cause of the problem can further be the technical structure of the underlying algorithm in an AI system. The structure, the defined target and the relevant variables included in the cycle of an algorithm can all play a role in this context, usually unintentionally, but intentional discrimination is also imaginable.

[151] See III.2.1.
[152] See Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 691ff; Phillip Hacker, 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law' (2018) 55 Common Mark. Law Rev. 1143, 1148f.
[153] See Frederik Zuiderveen Borgesius, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making* (Council of Europe 2018) p. 21.

2.1 Feature Selection

Initially, the definition of the target variables in an algorithm can contribute to algorithmic discrimination.[154] The target is the wanted outcome a machine learning algorithm should predict or for which it should create an optimization. When the required target leaves too much room for evaluation, it can happen that the algorithm combines variables that statistically lead to the goal but have a discriminatory effect.[155]

As a further cause is the structure of the algorithm, also referred to as feature selection.[156] The term feature selection describes the process of identifying and selecting the most important variables or attributes from a data set to create a predictive model. This step is crucial in machine learning as the selected features can have a significant impact on the performance, results and accuracy of the model.[157] There are three possible ways in which the structure of an algorithm can be causal in this context. First, the direct consideration of membership of a protected group, e.g. when explicitly taking gender into account; second, the consideration of an insufficient number of factors to evaluate members of a protected group with the same degree of accuracy as non-members, e.g., if fewer women have been hired to date, then data on women may be less reliable than data on male applicants; and third, using characteristics that serve as a proxy for protected characteristics[158], e.g. the average length of employment of all women is statistically reduced by women who pause due to childcare, thereby this indicator becomes a proxy for gender in hiring decisions.[159]

[154] Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 677.
[155] Ibid., 677ff.
[156] Joshua A. Kroll, Joanna Huey, et al., 'Accountable Algorithms' (2017) 165 U. Pa. L. Rev. 633, 681.
[157] See Solon Barocas and Andrew D. Selbst, 'Big Data's Disparate Impact' (2016) 104 Calif. Law Rev. 671, 688.
[158] On this problem see III.2.1.
[159] Joshua A. Kroll, Joanna Huey, et al., 'Accountable Algorithms' (2017) 165 U. Pa. L. Rev. 633, 681.

2.2 Intentional discrimination

A factor that is also conceivable is intentional discrimination. This cause is often described with the appropriate term of “masking”.[160] This refers to the intentional use of neutral or seemingly unrelated characteristics to mask the influence of protected characteristics such as race or gender in decision-making processes.[161] As mentioned earlier, the decision-making steps are not easy to understand and reconstruct, making it easy for developers to create an intentionally discriminatory application and hide behind the technology.

3. Lack of Diversity in Data Science

The creators and developers of AI systems are predominantly men which is also noteworthy, particularly with regard to gender bias.[162] This lack of diversity in the teams developing AI systems can result in a narrow understanding of fairness and inclusivity, inadvertently perpetuating biases in automated decisions. These structural inequalities in the AI sector will have an impact on the design, development and use of AI technologies and indeed are already doing so.[163] A mostly male work force can lead to unconscious biases being incorporated into algorithmic models, as male perspectives and experiences are likely to dominate the development process. This can result in algorithms not taking into consideration the diverse needs and experiences of women or other underrepresented groups, perpetuating gender and other forms of bias in AI

[160] See Solon Barocas and Andrew D. Selbst, ‘Big Data’s Disparate Impact’ (2016) 104 Calif. Law Rev. 671, 688; Joshua A. Kroll, Joanna Huey, et al., ‘Accountable Algorithms’ (2017) 165 U. Pa. L. Rev. 633, 682.
[161] Phillip Hacker, ‘Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law’ (2018) 55 Common Mark. Law Rev. 1143, 1149f.
[162] See Catherine D’Ignazio and Lauren F. Klein, *Data Feminism* (MIT Press 2020) p. 27ff.
[163] See Sarah Myers West, ‘Discriminating Systems: Gender, Race, and Power in AI’ (*AI Now Institute* 1 April 2019) <ainowinstitute.org> accessed 2 Dec 2024.

decision-making.[164] In addition, the lack of diversity on data science teams will limit critical discussions about equality and reduce the likelihood that biased outcomes will be recognized and addressed.

4. Interim findings

AI-based discrimination is a significant and probably much broader problem than currently recognized due to the opacity of many systems. Nevertheless, numerous cases already show that algorithms can disproportionately disadvantage marginalized groups in critical areas. This bias has both technical reasons and non-technical reasons, such as societal inequalities embedded in the decision-making processes. The question now is to determine whether the European Union's legal framework, and in particular the GDPR and the AI Act, addresses these concerns, and whether these provisions take full account of the complexity and subtlety of AI-induced discrimination.

V. Artificial Intelligence under the EU legal framework - Introduction

In order to present this comprehensively, firstly the European data protection law, the AI Act and their interplay will be presented in general terms, followed by an analysis of the two legislations in terms of their protection against discrimination by AI systems.

1. Data Protection Law and Artificial Intelligence

As seen above, data plays the key role within the scope of discriminatory AI decisions, which is why a look at the data protection law seems appropriate. In terms of data

[164] Janneke Gerards and Raphaële Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021), p. 90.

protection law, attention needs to be paid to the GDPR, which may offer legal options to counter this problem. The GDPR was adopted by the European Union on April 27th, 2016, and is applied since May 25th, 2018.[165] It sets out guidelines for the collection and processing of personal data of individuals within the European Union and strengthens their rights. Further, it also addresses the transfer of personal data outside of the EU. A reference to the right to data protection is also made in Art. 8 CFREU and Art. 16 TFEU, as well as in the national provisions of the member states[166], which indicates the importance of this right.

As has been the case since the beginning of the development of European data protection law[167], it can be assumed that the GDPR has a dual motivation: on the one hand, strengthen the fundamental right to the protection of personal data and, on the other hand, an economic objective, in particular the free movement of personal data within the EU.[168] These objectives become apparent in Art. 1 GDPR, as it mentions both the protection and rights of individuals with regard to personal data, Art. 1 (2) GDPR and the freedom of movement of personal data, Art. 1 (3) GDPR and can therefore be seen as a balance provision.[169] At this point and with a view to Recital 1, which establishes data protection as a fundamental right, evidently the protection of individuals is at the center of the GDPR. Based on this, some argue the protection of

[165] See Art. 99 (2) GDPR.
[166] For example, in Germany this is laid down in the constitution as the right to informational autonomy (Recht auf informationelle Selbstbestimmung), Art. 2 (1) in connection with Art. 1 (1) Grundgesetz.
[167] For an overview of the development see, Thomas Streinz, 'The Evolution of European Data Law' in Paul Craig and Gráinne de Búrca (eds), *The Evolution of EU Law* (3rd edn Oxford University Press 2021) p. 902ff; see also Spiros Simitis, Gerrit Hornung, et al., 'History and Motives of Data Protection Legislation' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) p. 2-19.
[168] Eugen Ehmann and Martin Selmayr, 'Introduction: Hintergrund und Entstehung der DS-GVO' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 10.
[169] Gerrit Hornung and Indra Spiecker gen. Döhmann, 'Art. 1: Subject matter and objectives' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 23.

personal data shall be the primary objective and priority of the GDPR.[170] Therefore, the core element of the GDPR is the required justification for the processing of personal data; in addition, it contains various rights for the data subject (Art. 12-23 GDPR) while imposing numerous obligations on the processor. It is therefore based on several fundamental principles, including lawfulness, fairness and transparency, purpose limitation, data minimization and accuracy, Art. 6 (1) (a)-(d) GDPR. In the present context, Art. 22 GDPR, which regulates ADM, is of particular importance. This provision prohibits automated decision-making if it produces legal or similarly significant effects, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention.[171] Thereby, it constitutes one of the most relevant GDPR rights for the individual. When considering these objectives, a tension is evident between the use of AI systems which are based on enormous amounts of data and the GDPR. Furthermore, a different objective and effect of data protection law compared to classical non-discrimination law becomes apparent. While non-discrimination law focuses on de facto unequal treatment and aims to protect affected individuals of discriminatory decisions, data protection law focuses on prior steps and protects at the processing stage.[172] In comparison, data protection law is therefore more preventive than protective in the aftermath.[173] For an introductory overview, a look is needed at the scope of application of the GDPR.

---

[170] Hielke Hijmans, 'Art. 1: Subject-matter and objectives' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 56.

[171] See Rec. 71 Sentence 1 GDPR.

[172] Mario Martini, *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz* (Springer 2019) p. 79.

[173] Gerrit Hornung and Indra Spiecker gen. Döhmann, 'Art. 1: Subject matter and objectives' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 4.
On the complex interplay, though not the core of this thesis, between non-discrimination law and data protection law, see Phillip Hacker, 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law' (2018) 55 Common Mark. Law Rev. 1143, 1143ff; see also Raphaël Gellert, Katja de Vries, et al., 'A Comparative Analysis of Anti-Discrimination

Initially, it should be mentioned that the term Artificial Intelligence or related terminology is not mentioned in the GDPR. However, this does not preclude the application. The material scope of application is set out in Art. 2 GDPR. According to Art. 2 (1) GDPR, it depends on whether, firstly, personal data[174] is involved and, secondly, whether this data is processed automatically wholly or partly.[175] The use of personal data is in principle conceivable[176]; furthermore, there are no restrictions on certain technical applications processing data automatically. Accordingly, this applies to any form of automated processing of personal data, including the deployment of AI systems that collect, analyze or make decisions based on such data. This also becomes clear with a view to Recital 15, which refers to technical neutrality and denies dependency on certain technologies.[177] The GDPR is therefore applicable to AI systems. The territorial scope of Art. 3 GDPR takes account of the fact that data processing is nowadays not bound by national borders.[178] Under Art. 3 (1) GDPR, it applies to the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not; according to Art. 3 (2) GDPR it even applies in certain cases to the processing by a controller or processor not established in the Union.

This thesis aims to examine the extent to which data protection law is able to help prevent discrimination by AI systems, therefore raising the question of the extent to

---

and Data Protection Legislations' in Bart Custers, Toon Calders, et al. (eds), *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases* (Springer 2013) p. 61-89.

[174] The Term 'personal data' is defined in Art. 4 (1) GDPR as follows, 'personal data' means any information relating to an identified or identifiable natural person ('data subject').

[175] The Term 'processing' is defined in Art. 4 (2) GDPR as follows, 'processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'.

[176] More on this, also with reference to the use cases presented, see VI.1.

[177] See Rec. 15 Sentence 1 GDPR.

[178] Stefan Ernst, 'Art. 3 Räumlicher Anwendungsbereich' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 1.

which data protection law may contribute to such discrimination. The GDPR may unintentionally contribute to discrimination by restricting access to personal data that could help mitigate discrimination.[179] There is an emphasis on data minimization, meaning only necessary data should be collected and processed, Art. 5 (1) (c) GDPR. While this is important for the protection of personal data, it can significantly limit the diversity of data and thus foster the reasons for discrimination mentioned above: over- and underrepresentation of certain groups in the data. Data protection law can further limit the extent to which AI models can be audited for discrimination as a measure to counteract discrimination. To assess the fairness of AI, access to sensitive attributes (e.g. ethnicity, gender) is often needed to measure potential bias.[180] However, data protection laws restrict the processing of such attributes, making it difficult to assess and address discrimination in AI outcomes.[181] In this respect, data protection law can in turn be regarded as a cause or intensifier of the problem at hand.[182]

Within the context of AI and data protection, data protection through AI can also be briefly mentioned. This includes AI being used to improve data protection and security, e.g. by identifying vulnerabilities, detecting unauthorized access and automatically complying with data protection regulations. By automating these processes, AI has the potential to reduce human error and shorten response times to potential data breaches, encouraging a stronger data protection approach.[183]

---

[179] See Zanda Davida and Dominik Lubasz, 'Privacy by Design – Searching for the Balance Between Privacy, Personal Data Protection and Development of Artificial Intelligence Systems' in Dariusz Szostek and Mariusz Załucki (eds), *Internet and New Technologies Law: Perspectives and Challenges* (Nomos 2021) p. 337, 345ff.

[180] See Marvin van Bekkum and Frederik Zuiderveen Borgesius, 'Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?' (2022) 48 Comput. L. & Secur. Rev. 105770, 3.

[181] Ibid., 5.

[182] Attention needs to be drawn here to the new provisions of the AI Act, in particular to Art. 10 AI Act, see VII.2.1.

[183] See Jan Geert Meents, 'Datenschutz durch KI' in Markus Kaulartz and Tom Braegelmann, *Rechtshandbuch Artificial Intelligence and Machine Learning* (C.H. Beck, Vahlen 2020) p. 445 ff.

2. Artificial Intelligence Act

In addition to data protection law, the recently adopted AI Act will be analyzed and set in relation to the GDPR. The European Commission first proposed the AI Act on April 21st, 2021[184] and besides this also introduced a revised coordinated plan on AI[185] aiming for the EU to be a "world-class hub for AI, while ensuring that AI is human-centric and trustworthy".[186] More than four years later, the legislation was formally adopted on May 21st, 2024 and since then establishes the first comprehensive legal framework on AI.[187] Some are expecting and assuming the AI Act to become a model for the global governance of AI, just as the GDPR has influenced the global protection of data.[188] The provisions will enter into force on August 2nd, 2026 at the latest, some of them earlier, as set out in Art. 113 AI Act.

The process of establishing a uniform regulation began back in 2018, when the EC recognized the need for a comprehensive AI regulatory framework and therefore released the European AI Strategy in April 2018.[189] A further landmark was the EU White Paper on AI.[190] The White Paper already proposed a regulatory framework for

[184] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts' COM (2021) 206.
[185] European Commission, 'Communication from the Commission to the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Fostering a European approach to Artificial Intelligence' COM (2021) 205.
[186] Ibid, p. 1.
[187] For a comprehensive overview of the European developments in the field of AI, see Nikos Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act* (Springer 2023) p. 23ff.
[188] Ceyhun Necati Pehlivan, 'The EU Artificial Intelligence (AI) Act: An Introduction' (2024) 5 Global Privacy Law Review 31, 33; in general on this thought see Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020).
[189] European Commission, 'Communication from the Commission to the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe' COM (2018) 237.
[190] European Commission, 'White Paper on Artificial Intelligence - A European approach to excellence and trust' COM (2020) 65 final.

safe and trustworthy AI, emphasizing a risk-based approach within the regulation which shall create an “ecosystem of excellence” and an “ecosystem of trust”.[191]

Accordingly, the AI Act pursues two objectives: on the one hand, the benefits offered by AI should be fully exploited, and on the other hand, the risks associated with AI are recognized and are to be prevented. This double-edged nature is already reflected clearly in the Recitals, as Recital 4 states that AI is a “fast evolving family of technologies that contributes to a wide array of economic, environmental and societal benefits across the entire spectrum of industries and social activities”.[192] Whereas Recital 5 states that “At the same time, depending on the circumstances regarding its specific application, use, and level of technological development, AI may generate risks and cause harm to public interests and fundamental rights that are protected by Union law”.[193] This double-edged objective is inherently contradictory, as more regulation to protect fundamental rights and democracy logically restricts innovation on the other hand. Regulation to protect the user is necessary and sensible in any case; however, to avoid an over-regulation, the objective of promoting innovation must be considered, as it has been done. Overall, the AI Act is intended to pursue a “human centric”[194] regulatory framework by applying a risk-based approach.

2.1 Scope of application

For a precise assessment, it is initially necessary to look at the scope of application of the AI Act. According to Art. 2 (1) AI Act it applies to providers and deployers operating

[191] Ibid, p. 3.
[192] See Rec. 4 Sentence 1 AI Act.
[193] See Rec. 5 Sentence 1 AI Act.
[194] The term human centric was already used in the ‘Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts’ COM (2021) 206, p. 1; see also Art. 1 AI Act and Rec. 1 Sentence 1 AI Act.

within the EU regardless of whether they are established in the Union or in a third country, which means the legislation has an extraterritorial scope of application.[195] This includes both public and private sector bodies developing and deploying AI systems. The extraterritorial approach is comparable to the GDPR which, as mentioned above, also applies to controllers or processors that are not established in the EU according to Art. 3 (2) GDPR. The following paragraphs of Art. 2 AI Act define, among other things, a number of exceptions to the applicability, e.g. according to Art. 2 (12) AI Act the exception with regard to free and open-source models or according to Art. 2 (8) AI Act the exception for research, testing and development activities on AI systems and AI models.

The material scope of application and thus the question of the definition of AI has already been addressed above.[196] It should be noted, however, that the regulation applies to AI systems (Art. 3 (1) AI Act) and AI models (GPAI in the meaning of Art. 3 (63) AI Act) which have different meanings in the context of the Regulation.[197] Art. 3 (1) AI Act defines the central term 'AI system' within the meaning of the Regulation and, as already mentioned, does so in an appropriate manner; on the one hand precisely enough for classification and on the other hand wide enough to include future technologies.[198] Even though it is indeed quite broad, this definition is not intended to be as precise as possible in the sense of a computer science approach.[199] This is why the definition used in the Regulation is connected to key characteristics, such as the capability to infer, in order to distinguish these systems from "simpler traditional

[195] Ceyhun Necati Pehlivan, 'The EU Artificial Intelligence (AI) Act: An Introduction' (2024) 5 Global Privacy Law Review 31, 33.
[196] See I.1.2.
[197] See Rec. 97 AI Act.
[198] See Rec. 12 Sentence 1 AI Act.
[199] Timon-Johannes Engel, 'Die KI-Verordnung - ein systematischer Überblick' (2024) 7-8 Kommunikation und Recht 445ff.

software systems or programming approaches".[200] In addition, the definition used is based on the one provided by the OECD, and thus creates a uniform framework.[201]

2.2 Risk-based approach

The regulation addresses and regulates both the development and the sale as well as the use and thus covers all aspects within the life cycle of an AI system.[202] Nevertheless, the AI Act is not generally viewed as a comprehensive regulation of AI, as the obligations of the AI Act are concretized at the level of directives or CE marking.[203] The key point is to design and develop AI systems in a safe way even before they are used, so the main focus of the regulation is on product safety.[204] Therefore it uses, as mentioned, a horizontal and "clearly defined" risk-based approach to regulate AI systems based on their potential harm to individuals and society.[205] In the area of technical regulation, the EU has increasingly turned to risk-based approaches, so this is not a new legal nature.[206]

This means AI systems are categorized in four categories, as unacceptable, high, limited or minimal risk. Unacceptable risk systems are prohibited according to Art. 5 AI

---

[200] See Rec. 12 Sentence 2 and 3 AI Act.
[201] OECD defines an AI system as 'a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.', see 'Recommendation of the Council on Artificial Intelligence' (*OECD* 3 May 2024) <legalinstruments.oecd.org> accessed 2 Dec 2024.
[202] Nikos Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act* (Springer 2023) p. 339.
[203] See Christoph Krönke, 'Das europäische KI-Gesetz: Eine Verordnung mit Licht und Schatten' (2024) 8 NVwZ 529, 530; see also Sonja Dürager, 'Highlights und Pain Points aus dem "KI-Gesetz" (Teil I)' (2024) 513 ecolex 898, 899.
[204] See Christoph Krönke, 'Das europäische KI-Gesetz: Eine Verordnung mit Licht und Schatten' (2024) 8 NVwZ 529, 530.
[205] Rec. 26 Sentence 1 AI Act.
[206] At the latest since the adoption of the European Digital Single Market Strategy; European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A Digital Single Market Strategy for Europe' COM (2015) 192 final; see also Giovanni De Gregorio and Pietro Dunn, 'The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age' (2022) 59 Common Mark. Law Rev. 473, 476.

Act because of the major risk they pose to the rights of individuals and democracy. Art. 5 (1) AI Act lists the concerned AI systems which are regarded as unacceptable risk, including social scoring systems, systems for biometric categorization or systems for assessing the risk of a natural person committing a criminal offense. This list is exhaustive for the time being and does not provide for the possibility of extension, meaning that an amendment to the legislation would be required to extend it to new prohibited applications, which could be slightly impractical given the rapid development of technology.[207] On the other hand, a prohibition represents the most significant interference in economic activity and must therefore be unambiguous and clear without room for interpretation. However, it should be mentioned that the list is not generally exhaustive, i.e. practices that are prohibited anyway, for example under data protection law or non-discrimination law, remain prohibited when used with AI.[208]

The main part of the regulation concerns high-risk applications with the strictest regulations and most requirements for providers. Art. 6 (1) and (2) AI Act set the requirements for classification as a high-risk application. According to Art 6 (1) AI Act, a high-risk application applies if it is used as a safety component, or a product covered by EU legislation listed in Annex I and is subject to third party conformity assessment in accordance with that Annex I legislation.[209] In this respect, it is part of the New Legislative Framework for the harmonization of product safety law.[210] These products, which fall under the EU harmonization regulations, include for example machines, toys, medical devices or cableway installations.[211] Likewise, a high-risk system exists if it is

[207] Kalojan Hoffmeister, 'The Dawn of Regulated AI: Analyzing the European AI Act and its Global Impact' (2024) 2 ZEuS 182, 196.
[208] Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 7.
[209] See Annex I AI Act.
[210] See Rec. 9 AI Act; See also 'New legislative framework' (European Commission) <single-market-economy.ec.europe.eu> accessed 2 Dec 2024.
[211] See Annex I No. 1, 2, 8, 11 AI Act.

listed Annex III.[212] These especially named high-risk applications are biometric identification, critical infrastructure, education, employment, access to public benefits and services, law enforcement, migration, asylum and border control and the administration of justice and democratic processes. The systems mentioned must further entail a considerable risk of harm to protected legal interests pursuant to Art. 6 (3) AI Act, namely health, safety or fundamental rights of natural persons. Of particular importance for the application of AI in decision-making is Art. 6 (3) AI Act, which gives the example of systems that only insignificantly influence the outcome of a decision-making process and are therefore not considered high-risk. If a high-risk application exists, the requirements set out in Art. 8-15 AI Act must be met. These include, among others, according to Art. 9 AI Act that a risk management system must be established, Art. 10 AI Act sets requirements for the data sets used, Art. 13 AI Act concerns transparency and Art. 14 AI Act requires the possibility of human oversight. Art. 16-27 AI Act establish obligations of the providers and deployers of high-risk systems. Key elements are the conformity assessment (Art. 16 (f) AI Act) in accordance with Art. 43 AI Act and the implementation of a quality management system, Art. 17 AI Act. Under Art. 27 AI Act, public bodies or private bodies that provide public services must carry out a Fundamental Rights Impact Assessment (FRIA).

Limited risk systems are subject to transparency obligations, e.g. it must be ensured that the user is aware of interacting with an AI if the systems are intended for this purpose, Art. 50 (1) AI Act. For applications with minimal risk, there are no specific regulations, they can be used flexibly for everyday use. An example of a minimal risk

[212] See Annex III AI Act.

application are simple spam filters.[213] Art. 95 AI Act only stipulates voluntary compliance for these applications.

The risk-based approach of the regulation can therefore be illustrated as follows:

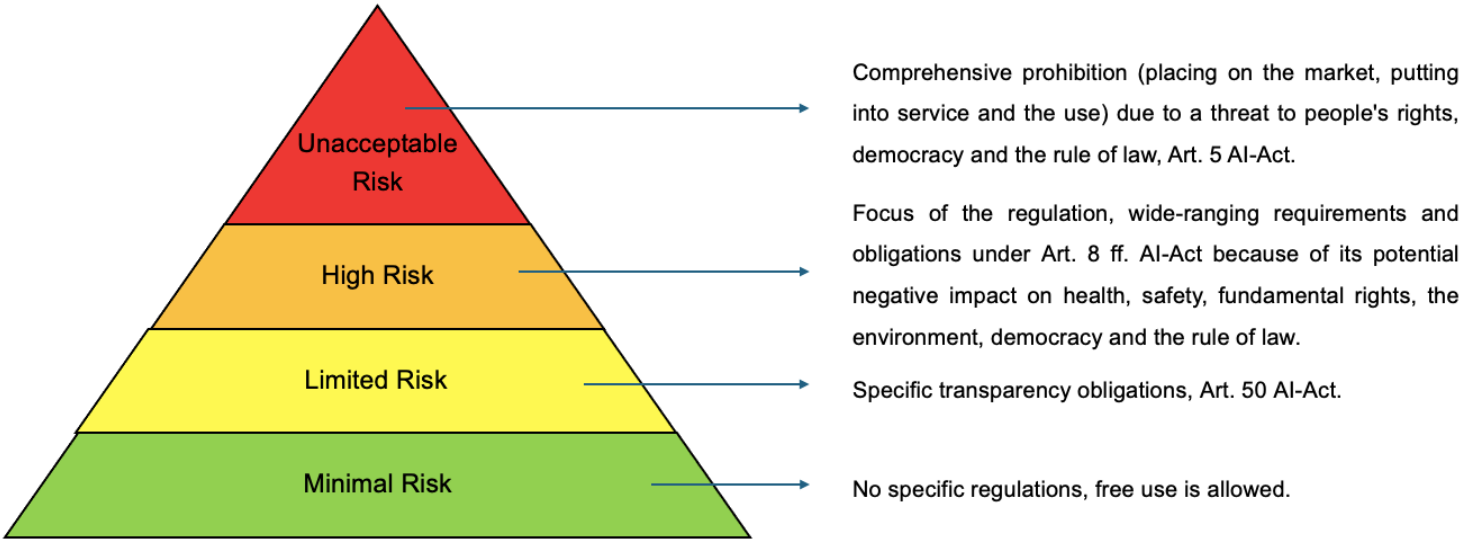

Illustration 2: The risk-based approach of the AI-Act.[214]

Irrespective of the four risk categories, regulations also exist for GPAI models according to Art. 51-56 AI Act. It is clearly stated that these models are to be distinguished from the concept of AI systems.[215] GPAI includes systems that are highly generalized and capable of competently performing a wide range of different tasks, regardless of how the model is brought to market, and that can be integrated into a variety of downstream systems or applications, Art. 3 (63) AI Act. For example, this includes models such as GPT-4 from OpenAI or Llama 3.1 from Meta, as well as other

[213] Marcus von Welser, 'Die KI-Verordnung – ein Überblick über das weltweit erste Regelwerk für künstliche Intelligenz' (2024) 15 GRUR-Prax 485, para. 33.
[214] Based on the visualization of the risk-based approach by the EC, see 'Shaping Europe's digital future: AI Act' (*European Commission* 14 Oct 2024) <digital-strategy.ec.europa.eu> accessed 2 Dec 2024.
[215] See Rec. 97 Sentence 1 AI Act.

Large Language Models.[216] Depending on the degree of risk, the provider is subject to various obligations, such as documentation or risk assessment obligations.

2.3 Innovation, individual's rights and penalties

In addition to the extensive requirements and obligations, Art. 57-63 AI Act contain measures in support of innovation to realize the second objective. Furthermore, another important element of the regulation are the rights of the affected persons set out in Art. 85ff. AI Act. According to Art. 85 AI Act every person has the right to submit complaints about infringements of the provisions of this regulation. Furthermore, any person who has been subject to a decision based on a high-risk system and which produces legal effects or similarly significantly affects has the right under Art. 86 AI Act to obtain information and a meaningful explanation of the precise way in which the system was used in the decision-making process. This right can be regarded as the counterpart to Art. 22 GDPR, that prohibits being subject to a fully automated decision.[217] Furthermore, Art. 99ff. AI Act stipulates penalties for infringements. These can be up to EUR 35 000 000 or, if the offender is an undertaking, up to 7 % of its annual worldwide turnover. This level of possible penalties contributes to the significance of the regulation.

3. Interplay between GDPR and AI Act

The question arises to what extent these two EU laws are connected and interplay with each other. Art. 2 (7) AI Act clarifies that the protection of personal data and the GDPR are not affected by the AI Act, the obligations of the processors continue to apply, and

[216] Marcus von Welser, 'Die KI-Verordnung – ein Überblick über das weltweit erste Regelwerk für künstliche Intelligenz' (2024) 15 GRUR-Prax 485, para. 34.
[217] Timon-Johannes Engel, 'Die KI-Verordnung - ein systematischer Überblick' (2024) 7-8 Kommunikation und Recht 445ff.

the data subjects still enjoy all the rights of the GDPR.[218] Both initially represent a form of legal regulation for technical risks and challenges, nonetheless, a few key differences stand out. The regulatory nature of the two legislations differs fundamentally: while the GDPR follows a rights-based approach by providing individual data subjects with various rights, the AI Act follows a risk-based approach by imposing requirements and obligations depending on the degree of risk.[219] In contrast to the GDPR, the AI Act does not focus on the rights of individuals per se, but on regulatory standards for the development and use of AI. These standards ensure security, accountability and transparency, especially for high-risk systems. Nevertheless, in part, both laws are understandably classified as risk-based and the differentiation is rather made in the application of this approach: the GDPR follows a "bottom-up approach" and the AI Act a "top-down approach".[220] The AI Act restricts discretion and provides the risk-based system instead of entrusting the providers and users of AI systems to develop their own risk mitigation system, as it is the case in the GDPR.[221] However, this fundamental distinction in the nature of the regulation is not the only difference; the material scope of application also differs as seen above. The scope of application for one legislation may be opened while the other legislation does not apply.[222] The GDPR applies as soon as personal data is processed, which does not necessarily have to be related to AI, meaning that the AI Act would not apply. Conversely, AI systems do not necessarily process any personal data. However, a

[218] See Rec. 10 AI Act.
[219] Josephine Wolff, Willian Lehr, et al., 'Lessons from GDPR for AI Policymaking' (2024) 27 Va. J. L. & Tech. 1, 20.
[220] Giovanni De Gregorio and Pietro Dunn, 'The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age' (2022) 59 Common Mark. Law Rev. 473, 477ff.
[221] Ibid, 490.
[222] See Ceyhun Necati Pehlivan, 'The EU Artificial Intelligence (AI) Act: An Introduction' (2024) 5 Global Privacy Law Review 31, 40.

wide area includes cases in which AI systems are trained with personal data, meaning that both regulations apply in a complementary manner.[223]

Besides these two fundamental differences, there are also parallels and an interplay. Such an interplay can be identified in relation to Art. 22 GDPR and the obligation under the AI Act to provide human oversight. Both regulations impose requirements on human oversight.[224] According to Art. 22 GDPR, individuals have the right not to be subject to solely automated decisions that significantly affect them. The AI Act, while not primarily, also aims to protect fundamental rights by requiring human oversight for high-risk AI systems. Art. 14 AI Act requires that high-risk AI systems must be equipped with human oversight capabilities to ensure that humans can intervene during the operation of the AI system. This level of human oversight could help to ensure that such AI systems do not fall under the GDPR's ADM provisions, as actual human intervention ensures that there is no longer a fully automated decision.[225]

Parallels are also visible in the requirements and compliance systems, as many of the risk management and governance obligations laid down in the AI Act are structurally similar to those of the GDPR.[226] The obligation to set up a risk management system pursuant to Art. 9 AI Act is conceptually similar to the data protection impact assessment (DPIA) under Art. 35 GDPR. Both, the risk management system and the DPIA are risk assessment tools to identify and evaluate the risks posed by AI systems on the one hand and the use of personal data on the other. The AI Act also provides

[223] Ibid.
[224] Tristan Radtke, 'Das Verhältnis von KI-VO und Art. 22 DS-GVO unter besonderer Berücksichtigung der Schutzzwecke' (2024) 8 RDi 353, 353.
[225] See James Clark, Muhammed Demircan, et al., 'Europe: The EU AI Act's relationship with data protection law: key takeaways' (*DLA Piper* 25 April 2024) with additional references <privacymatters.dlapiper.com> accessed 2 Dec 2024.
[226] Ceyhun Necati Pehlivan, 'The EU Artificial Intelligence (AI) Act: An Introduction' (2024) 5 Global Privacy Law Review 31, 40.

for further assessments to monitor compliance with the regulations and obligations. Firstly, the conformity assessment for high-risk AI systems under Art. 43 AI Act, as it is intended to ensure the requirements are met before they are placed on the EU market. This assessment tool includes a focus on fundamental rights as part of a broader review just as the second assessment tool which is specifically focused on fundamental rights.[227] According to Art. 27 AI Act, some operators of AI systems must create a FRIA before the AI system is permitted to be used; this includes a risk assessment of the extent to which fundamental rights of the individual are affected and the extent to which this can be counteracted. Although the DPIA and the FRIA both have the same core features, the scope of the FRIA assessment is different as it considers all fundamental rights, while the DPIA focuses only on data protection.[228]

To summarize, the GDPR and the AI Act form a multi-layered legal framework, each with a different scope and approach. The GDPR generally applies to the processing of personal data and focuses on data protection in all sectors, while the AI Act specifically regulates AI systems with an impact on health, safety or fundamental rights.

VI. GDPR as a safeguard against discrimination through AI systems

Data represents the fundamental basis for AI systems. For this reason, it will be examined to what extent the GDPR provides protection mechanisms against discrimination by AI systems or how far the GDPR can be used for such protection, by

[227] Alessandro Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template' (2024) 54 Comput. L. & Secur. Rev. 106020, 6.
[228] Ibid, 4.

first looking at the use of personal data in AI systems - with regard to the use cases - followed by a detailed look at relevant individual provisions.

1. Processing of personal data in AI based decision-making

As already explained, the GDPR applies to the processing of personal data by AI systems. However, specific determination must still be made to the extent to which AI systems used for decision-making carry out such processing and whether the use cases on which the examination is conducted contain personal data.

Processing is legally defined in Art. 4 (2) GDPR as follows: any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means. Further, the definition provides a number of use cases that represent such processing, like collection, organization or storage. Consequently, the concept of processing is very broad and therefore basically includes any form of data processing.[229] Given this assumption, an AI system used for ADM clearly involves 'processing' at various stages of the decision-making process. For example, data is collected, i.e. data about the data subject is obtained[230]; data is also stored, i.e. it is saved on a data carrier for future use[231]; the data is also organized and structured to create some kind of structure within the data in the AI systems.[232] Further required is the involvement of personal data. Personal data means any information relating to an identified or identifiable natural person, Art. 4 (1) GDPR. The ECJ interprets the term very widely, reflecting “the aim of the EU legislator to assign a wide scope to that

[229] Stefan Ernst, ‘Art. 2 Sachlicher Anwendungsbereich’ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 5.
[230] Stefan Ernst, ‘Art. 4 Begriffsbestimmungen’ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 23.
[231] Stefan Ernst, ‘Art. 2 Sachlicher Anwendungsbereich’ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 25.
[232] Ibid., para. 26.

concept, which is not restricted to information that is sensitive or private, but potentially encompasses all types of information".[233] Therefore, it includes objective and subjective information.[234] Information such as name, age, gender, zip code or date of birth and also, for example, financial situation, contractual relationships or consumer behavior are included if these are related to an identifiable person.[235] With regard to the use cases of AI in decision-making processes, the following therefore applies:

1.1 Application of AI systems in Law Enforcement and Border Control

For the application in risk assessment tools, like the COMPAS system in the US, the use of personal data is central both in training through the use of historical data and in deployment through the creation of predictions for individual defendants assessed on the basis of personal profiles. Personal data is likewise used in predictive policing. As mentioned above, location data, crime records and demographic information containing potentially identifiable data are used to predict where crime is likely to occur.

In this field, a look needs to be taken at the exceptions provided in Art. 2 (2) GDPR, and to Art. 2 (2) (d) GDPR. Accordingly, the GDPR does not apply to processing by the competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offenses or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security – meaning it does not apply if the Law Enforcement Directive (LED)[236] applies. This establishes an explicit

[233] Case C-434/16 *Peter Nowak v Data Protection Commissioner* [2017] ECLI:EU:C:2017:994, para. 34.
[234] Peter Gola, 'Art. 4 Begriffsbestimmungen' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 6.
[235] Stefan Ernst, 'Art. 4 Begriffsbestimmungen' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 14.
[236] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89; see Vagelis Papakonstantinou and Paul De

exclusion of “law-enforcement data processing” from the GDPR, meaning that data processing is generally permitted in these contexts, provided that the competent authority acts.[237] However, a precise distinction between the applicability of the GDPR and the Law Enforcement Directive can be difficult to make, as it is not only the competence of the processor that matters, since the GDPR applies if the purposes of the processing do not fall under the Law Enforcement Directive, as stated in Recital 19 GDPR.[238] Consequently, it always depends on the individual case and is to be determined according to the exact circumstances.

In the area of border control, personal data are used for profiling, including biometric data such as facial expressions to recognize emotions, passport information and travel history. Especially biometric identification and face recognition systems possess the ability to identify persons at a distance, based on historical images or videos.[239]

1.2 Application of AI systems in the Administrative Sector

In the area of administration, beginning with the labor market, personal data also has a crucial function. In the labor market sector, the use of classification systems is based on personal data, including gender, resumes, work histories and even potentially sensitive information like health data. Such systems need to be trained on historical personal data, plus personal data is also used to evaluate candidates. A further application concerns social welfare and fraud detection. Here, too, personal data is incorporated into both the training and the application, as AI systems analyze personal

Hert, ‘Art. 2: Material Scope’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023), para. 68.

[237] Orla Lynskey, ‘Criminal justice profiling and EU data protection law: precarious protection from predictive policing’ (2019) 15 International Journal of Law in Context 162, 164.

[238] Vagelis Papakonstantinou and Paul De Hert, ‘Art. 2: Material Scope’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023), para. 70ff.

[239] Costica Dumbrava, *Artificial Intelligence at EU borders: Overview of applications and key issues* (European Parliament 2021) p. 28.

data such as income records, social benefit history and demographic data to assign social benefits or detect social benefit fraud. The same applies to the use of AI systems in the admission and allocation of university places. Again, personal data is used by relying on academic data, demographic data such as age, gender, place of residence and socio-economic data such as education and income.

Thus, the use of AI systems for decision-making or decision support in these fields are within the scope of the GDPR. Of course, this also requires identifiability to a specific person, which is not always easy to determine. Importantly, the processor does not necessarily have to be able to carry out the identification itself, rather it depends on whether a third party is able to carry out such an identification, while taking into account all means that can reasonably be used.[240] However, in the present framework, which relates to decision-making systems, this problem does not need to be deepened, as it is assumed to be a situation in which the information can in any case be assigned to a specific identifiable person about whom the decision is to be made.[241]

1.3 Anonymized data as an exception

An exception to the application of the GDPR exists for anonymized data, as set out in Recital 26 Sentence 5. According to this, the GDPR does not apply to anonymous information, i.e. information that does not relate to an identified or identifiable natural person, or to personal data that has been anonymized in such a way that the data subject is not or no longer identifiable. Yet, this does not apply to pseudonymized data, personal data that has been pseudonymized is still subject to the provisions of the

[240] See Rec. 26 GDPR; Achim Klabunde and Anna Zsófia Horváth, 'Art. 4 Begriffsbestimmungen' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 17.
[241] On the problem of identifiability see Ibid, para. 10ff; see also Peter Gola, 'Art. 4 Begriffsbestimmungen' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 7ff.

GDPR.[242] Pseudonymization is merely a different form of storage, whereby the entire information content of the data is still available and the individual reference of each piece of information can be created without great difficulty.[243] The question arises as to whether the anonymization is an easy method for processors to circumvent the requirements of the GDPR. However, this is not only in the interest of the processor, but it also protects the individual who is subject to an automated AI based decision. One problem, though, lies in the fact that it is difficult or even impossible to anonymize big data, which is why all data in the case of big data can be regarded as personal data.[244] Although there are methods of de-identification for big data, through which data can no longer be assigned to individuals[245], there are also methods of re-identification, meaning anonymized data can often be identified again and assigned to an individual, and these technical possibilities are constantly improving.[246]

Regarding the issue to be examined here, however, no in-depth discussion is required. For sure, anonymization and even the pseudonymization of data could certainly contribute to protection against discrimination. As anonymization removes identifying information from the data, which means that AI systems do not directly relate to protected characteristics such as ethnicity, gender or age. Likewise, by replacing identifiable information with non-genuine identifiers, pseudonymization may help to ensure that decisions are not influenced by personal identities. Meaning, if certain characteristics are removed, the AI algorithm cannot use them as a basis for its

[242] Luca Tosoni, 'Art. 4 (5) Pseudonymisation' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 134.
[243] Achim Klabunde and Anna Zsófia Horváth, 'Art. 4 Begriffsbestimmungen' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 19.
[244] See Volker Boehme-Neßler, 'Das Ende der Anonymität: Wie Big Data das Datenschutzrecht verändert' (2016) 7 Datenschutz und Datensicherheit, 419ff.
[245] Omer Tene and Jules Polonetsky, 'Privacy in the Age of Big Data' (2012) 64 Stan. L. Rev. Online 63, 65 <stanfordlawreview.org> accessed 3 Dec 2024.
[246] See on this already in 2010, Paul Ohm, 'Broken Promises of Privacy: responding to the surprising failure of anonymization' (2010) 57 UCLA L. Rev. 1701ff.

decision and discriminate on these grounds. Yet, even with anonymized data an AI is likely to recognize correlations, patterns and connect data points and thus still assign people to certain groups. AI and in general methods of computational statistics increase the identifiability of seemingly anonymous data, as they make it possible to link unidentified data (including anonymized or pseudonymized data) to the individuals concerned.[247] In addition, certain information is necessary for a fair basis for decision-making; in the present thesis, the focus is on personal decisions that must be based on personal information. The quality of the data set on which the AI is based tends to decrease through anonymization because details are removed from the set.[248] Which is why the relevance for discrimination in this context is rather limited.

2. Principles of data processing, Art. 5-11 GDPR

Having established that the GDPR is applicable and plays a relevant role, it remains to be examined to what extent the principles for processing of personal data laid down in Art. 5-11 GDPR can contribute to combating discrimination by AI systems and whether AI systems can comply with these principles at all or whether this hinders the use of AI systems altogether. These key principles are lawfulness, fairness and transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity and confidentiality as well as accountability, Art. 5 (1), (2) GDPR. Art. 5 GDPR is therefore the guiding provision for all following regulations of the GDPR, as all further provisions aim to implement these principles precisely.[249] The principles are binding - the

[247] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 36.
[248] Felix Bauer, Stefan Buchberger, et al., 'Machine Learning und die Transparenzanforderungen der DS-GVO' (*bitkom* 2018) p. 30 <bitkom.org> accessed 3 Dec 2024; see also Tina Gausling, 'KI und DS-GVO im Spannungsverhältnis' in Johannes Graf Ballestrem, Ulrike Bär, et al. (eds), *Künstliche Intelligenz: Rechtsgrundlagen und Strategien in der Praxis* (Springer 2020) p. 19.
[249] Alexander Roßnagel and Philipp Richter, 'Art. 5: Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 7.

relevance of this is particularly emphasized by the sanction provision of Art. 83 (5) (a) GDPR, which provides a fine for violations.[250] Although a sanction seems unlikely due to the requirement of legal certainty, since they are rather general and abstract requirements and need to be concretized for application, this is already achieved by the remaining provisions of the GDPR.[251] Para. 2 of Art. 5 GDPR explicitly states the addressee, namely the controller who is responsible for compliance with the principles. The controller is, according to Art. 4 (7) GDPR, any "natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".

2.1 Lawfulness, fairness and transparency, Art. 5 (1) (a) GDPR

Art. 5 (1) (a) GDPR stipulates that personal data needs to be processed "lawfully, fairly and in a transparent manner in relation to the data subject". Regarding the lawfulness principle, every processing of personal data requires either consent of the data subject concerned or another legitimate basis as set out in Art. 6 (1) GDPR.[252] The list of these six legal bases is final and exhaustive.[253] Obviously, this requirement serves the interests of every person affected by an AI-based decision, as the processing of personal data always requires a legal basis. A more detailed look will only be taken at the legal bases that are particularly relevant and the potential problems that can arise.

[250] Eike Michael Frenzel, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 1ff.
[251] Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 16.
[252] See Rec. 40 GDPR.
[253] Waltraut Kotschy, 'Art. 6 Lawfulness of processing' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 329.

Conceivable here is firstly the consent of the data subject pursuant to Art. 6 (1) (a) GDPR and specified in Art. 7 and 8 GDPR. Consent is defined in Art. 4 (11) GDPR as any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her. Of particular importance is the free nature as well as the necessary comprehensive knowledge regarding the data processing. At a minimum, the data subject concerned needs to be informed of the identity of the controller and the purposes of the processing for which the personal data are intended.[254] Furthermore, the data subject must have had a genuine choice since consent should not be considered as freely given if the data subject does not have a free choice or is not in a position to refuse or withdraw consent without detriment.[255] With regard to consent under the GDPR for the processing of personal data in AI applications, the following problems arise. Firstly, for consent to be effective, data subjects must be informed, i.e. being able to clearly understand what they are consenting to. When it comes to AI, this can be a challenge as the data processing intelligence is often complex and involves algorithms that are potentially not easy to understand. Essentially, AI systems that use personal data must be explainable to a certain extent. However, as it has been seen before AI's operations cannot be explained in every case because of the black box issue. This means consent is often neither based on real knowledge nor on a real possibility to choose because of a lack of awareness.[256] Besides, the general impracticability of providing informed and reasoned consent in each individual case of processing in the current socio-

[254] See Rec. 42 Sentence 4 GDPR.
[255] See Rec. 42 Sentence 5 GDPR.
[256] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 41.

technical conditions is an additional problem.[257] Personal data is processed in numerous cases where consent is required so that it is not possible for the individual to keep track of them all, i.e. a rational assessment of the potential benefits and risks in all these cases is likely impossible.[258] Secondly, the freedom to decide as having a real choice can be problematic. This freedom is lacking if there is an imbalance of power, which is particularly the case if the controller is a public authority, because it is unlikely that the consent was given freely in all the circumstances of that specific situation.[259] In the given examples of an AI-based decision by public authorities, such an imbalance of power may therefore exist. As a consequence, the requirement of lawful processing carried out by public authorities laid down in Art. 6 (1) (e) GDPR generally cannot be substituted by individual's consent.[260] Consent is therefore generally possible in cases of the use of AI for decision-making on the basis of personal data - albeit with the difficulties outlined above and always aware of the possibility of withdrawal by the data subject (Art. 7 (3) GDPR) - however, this does not apply for the use by public authorities.

The already mentioned and further legal bases of Art. 6 (1) GDPR relevant in this context, is the processing for the exercise of a task carried out in the public interest or in the exercise of official authority vested in the controller, Art. 6 (1) (e) GDPR. This legal basis can be considered as the general basis for processing in the public sector.[261] National, regional and municipal governments and their authorities to which

[257] Giovanni Sartor, 'Art. 6 Lawfulness of Processing' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 22.
[258] Ibid.
[259] See Rec. 43 Sentence 1 GDPR.
[260] Waltraut Kotschy, 'Art. 6 Lawfulness of processing' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 330.
[261] Giovanni Sartor, 'Art. 6 Lawfulness of Processing' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 47.

tasks are specifically allocated are covered by this provision, as well as private bodies to which a task has been delegated.[262] However, it should be noted that this does not constitute a general legal authorization for the processing of personal data in the public interest[263]; rather, there must still be a legal basis under EU or Member State law.[264] Furthermore, it is questionable whether a public interest can be given if it does not comply with the legal system, such as if discriminatory results are achieved.[265] The further condition of necessity can be considered quite flexible, which means it is sufficient that another way of performing the task is less effective.[266] Besides, Art. 6 (1) (c) GDPR also comes into consideration, which permits processing for compliance with a legal obligation to which the controller is subject. This includes only legal obligations laid down in EU or Member State law (Art. 6 (3) GDPR), meaning that contractual obligations are not considered.[267] Public authorities can be permitted to process data on this legal basis insofar as the law imposes an obligation in this regard. It can therefore be concluded that data processing by public authorities using AI systems is mostly covered by Art. 6 (1) (c) or (e) GDPR as long as the necessity is given.

Furthermore, Art. 5 (1) GDPR establishes the principles of fairness and transparency. The former implies a behavior during the processing that generates trust, i.e. a

[262] Eike Michael Frenzel, 'Art. 6 Rechtmäßigkeit der Verarbeitung' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 23.
[263] See Rec. 45 Sentence 1 GDPR.
[264] Sebastian Schulz, 'Art. 6 Rechtmäßigkeit der Verarbeitung' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 51.
[265] On the same consideration in relation to the legitimate interest under Art. 6 (1) (f) GDPR, see Horst Heberlein, 'Art. 6 Rechtmäßigkeit der Verarbeitung' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 40.
[266] Giovanni Sartor, 'Art. 6 Lawfulness of Processing' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 49; see also, Case C-524/06 *Heinz Huber v Bundesrepublik Deutschland* [2008] ECR I 9725.
[267] Giovanni Sartor, 'Art. 6 Lawfulness of Processing' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 41.

behavior that is literally fair.[268] On the one hand, this includes “informational fairness” [269], which is linked to transparency, requiring the data subject to be informed about the existence of the processing and its purposes.[270] In this context, the problem mentioned above comes into play again - the difficulty of explaining the processing of personal data in AI systems. On the other hand, “substantive fairness”[271] is covered by the principle, which focuses on the outcomes of the processing and the effects it has on the data subjects.[272] This principle of fairness is closely related to the problem of AI-based decisions, since fairness is arguably lacking when the outcome is discriminatory. Accordingly, AI applications could already infringe the GDPR for this reason. Indeed, the broad scope of application of the principle together with the rather open structure results in a flexible provision that can be used to restrict automated decision-making systems with respect to fundamental rights.[273]

The principle of transparency of processing means the data subject must be provided with information regarding the collection, use, retrieval or other processing of personal data and the extent to which the personal data is or will be processed.[274] This

[268] Alexander Roßnagel and Philipp Richter, ‘Art. 5 Principles relating to processing of personal data’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 31.
[269] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 44; also referred to as ‘Procedural fairness’, see Andreas Häuselmann and Bart Custers, ‘Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR’ (2024) 52 Comput. L. & Secur. Rev. 105942, p. 3.
[270] See Rec. 60 Sentence 1 GDPR; independence of this principle is sometimes questioned and only seen in connection with transparency, see e.g. Alexander Roßnagel and Philipp Richter, ‘Art. 5 Principles relating to the processing of personal data’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 30; see also Horst Heberlein, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten‘ in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 16.
[271] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 45; see also Andreas Häuselmann and Bart Custers, ‘Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR’ (2024) 52 Comput. L. & Secur. Rev. 105942, p. 3.
[272] Andreas Häuselmann and Bart Custers, ‘Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR’ (2024) 52 Comput. L. & Secur. Rev. 105942, p. 4.
[273] Lee A. Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making' in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford, 2019 online edn) p. 260.
[274] See Rec. 39 Sentence 2 GDPR.

information must be easily accessible and understandable, with a clear and simple language being used.[275] Furthermore, the principle is strengthened by the rights of the data subject to information in accordance with Art. 13 and Art. 14 GDPR and the right of access under Art. 15 GDPR. With regard to AI, this transparency principle requires an adequate explanation of the functioning of the underlying algorithm.[276] This contains also "to provide plain and easy-to-understand information on the sources of data/input, factors, processes and/or logic that led to the prediction, content, recommendation or decision, to enable those affected by an AI system to understand the output".[277] Due to the opacity of many AI applications, this is difficult and often only barely achievable. Therefore, it is at least a fundamental contribution to the protection against discriminatory AI decisions, as the first and most important aspect to counteract is the understanding of the person concerned.

2.2 Purpose limitation, Art. 5 (1) (b) GDPR

The principle of purpose limitation ensures that personal data is only collected for specified, explicit and legitimate purposes and is not further processed in a way that is incompatible with these purposes. This principle is the central general principle which can be seen as the cornerstone of data protection and a prerequisite for most of the other key principles and requirements.[278] There is a connection between this principle and the principle of lawfulness, as Art. 6 (1) (a) GDPR requires consent for a specific

[275] See Rec. 39 Sentence 3 GDPR.
[276] Horst Heberlein, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 18 with additional reference.
[277] 'Recommendation of the Council on Artificial Intelligence' (*OECD* 3 May 2024) <legalinstruments.oecd.org> accessed 3 Dec 2024.
[278] Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 45; Cecile de Terwangne, 'Art. 5 Principles relating to processing of personal data' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 315.

purpose. When using AI in general and in particular when dealing with big data, a tension can be seen in relation to this principle, as these AI systems are designed to deliver a variety of results with the underlying data.[279] The intended function and use of big data are diametrically opposed to the principle of purpose limitation.[280] However, this is not overly important in the present context, assuming that the intended purpose - such as the assessment of employment opportunities or the verification of the granting of social benefits - is strictly adhered to and no other purpose follows.

2.3 Data minimization, Art. 5 (1) (c) GDPR

According to Art. 5 (1) (c) GDPR personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed. Personal data must therefore be limited to what is necessary for the purposes.[281] Logically, this again tends to contradict the concept of big data.[282] But this principle should also be considered in connection with proportionality, meaning any use of more personal data should not contradict this principle as long as it has advantages for the individual user or concerned person.[283] The reduction of use of personal data must therefore be as low as possible in relation to the purpose of processing.[284] As already mentioned, AI is based on data and the ability to produce effective outcomes depends on the underlying data. Since this generally offers advantages for those affected by a data-based AI

[279] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 45.
[280] Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 88.
[281] See Rec. 39 Sentence 7 GDPR.
[282] See Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 99.
[283] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 47.
[284] Eike Michael Frenzel, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten ' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 34.

decision at first glance, a proportion can be assumed here. However, this cannot apply if the data basis is biased and thus leads to discriminatory results, because the requirement to be necessary not only relates to the quantity but rather also to the quality of the personal data.[285] Consequently, the use of biased data as training data or input data may conflict with this principle.

2.4 Accuracy, Art. 5 (1) (d) GDPR

Personal data must further be accurate and, where necessary, kept up to date. Of course, this applies to AI systems that are intended to make decisions or provide support for decision-making; the input and underlying data of the data subjects must be correct. One problem associated is the potential opacity of the AI system used for decision-making, i.e. it is not possible to say for sure what data has been used and if it is correct or whether the self-learning system has evolved in such a way that the data is no longer accurate and/or correctly combined.[286]

2.5 Storage limitation, Integrity and confidentiality, Art. 5 (1) (e), (f) GDPR

Personal data shall further be kept in a form which permits identification of data subjects for no longer than necessary for the purposes for which the personal data is processed. Difficulties can also arise here: on the one hand, because the algorithms of a self-learning AI use the data even after the purpose has been completed to learn

[285] Cecile de Terwangne, 'Art. 5 Principles relating to processing of personal data' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 317.
[286] Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 119.

again on that basis and, on the other hand, it is not always possible to remove data, especially training data, from the AI model again.[287]

The final principle of the GDPR requires that personal data is processed in a manner that ensures appropriate security of it, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures. This serves to support and protect the other principles listed above and thus guarantees the right to data protection.[288]

2.6 Interim findings – data protection principles and AI

In conclusion, the principles of the GDPR set out in Art. 5 (1) are crucial to reduce the risk of discriminatory outcomes when AI is used for decision-making by public authorities. Compliance with the principles like fairness, transparency and accuracy will enable public authorities to ensure that AI systems do not perpetuate bias or unfairly prejudice individuals. Further purpose limitation and data minimization help to limit the use of AI to relevant data and thus reduce the risk of discriminatory or disproportionate decisions. Overall, the principles therefore also serve as protective measures in this context - AI and decision-making.

3. Sensitive Data, Art. 9 GDPR

Art. 9 GDPR establishes specific rules for the processing of special categories of personal data, also referred to as sensitive data[289], which by their nature are particularly sensitive in relation to fundamental rights and freedoms, since the context

[287] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 49.
[288] Alexander Roßnagel and Philipp Richter, 'Art. 5 Principles relating to processing of personal data' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 133.
[289] See Rec. 10 Sentence 5 GDPR.

of their processing could create significant risks to it.[290] Art. 9 (1) GDPR establishes a general prohibition to process sensitive data and contains an extensive list of sensitive data.[291] These include all personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. The term ‘revealing’ is of particular importance as it leads to the inclusion of data that is not directly related to the information about the person but reveals it in a certain way.[292] This means in principle, even simple data such as a name or a photo can allow inferences to be drawn about for example the ethnicity of a person, so that all data could be classified as sensitive. However, to avoid this, the category only includes data that is captured in order to derive special categories from it.[293] It is obvious that the above information about an individual carries a great potential for risk and, in particular, for discrimination which is why this provision is aimed at preventing discrimination.[294] Nevertheless, it should be noted that this list also leaves out relevant information that could entail such risks, such as the economic situation.[295]

[290] See Rec. 51 Sentence 1 GDPR.
[291] Ludmila Georgieva and Christopher Kuner, ‘Art. 9 Processing of special categories of personal data’ in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 373.
[292] András Jóri, ‘Art. 9 Processing of special categories of personal data’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 8.
[293] Ibid, para. 9.
[294] Eike Michael Frenzel, 'Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten‘ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 6 with additional reference; Alexander Schiff, 'Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten‘ in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 8.
[295] Eike Michael Frenzel, 'Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten‘ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 7.

Exceptions to the general prohibition of processing laid down in para. 1 are exhaustively laid down in Art. 9 (2) GDPR.[296] In the area of using AI systems for decision-making or support by public authorities, the following exceptions under Art. 9 (2) GDPR come into consideration: firstly, explicit consent can be given (Art. 9 (2) (a) GDPR), whereby this is characterized by more stringent requirements than, e.g. the consent of Art. 6 (1) (a) GDPR as it must have been expressly given[297]; secondly, if the processing is necessary to carry out obligations and exercising specific rights of the controller or the data subject in the field of employment, social security and social protection law (Art. 9 (2) (b) GDPR), which also includes providing and accounting of social benefits[298]; and thirdly, if a substantial public interest exists (Art. 9 (2) (g) GDPR), whereby this is again to be viewed as narrower than the public interest within the meaning of Art. 6 (1) (e) GDPR and requires a balance between the public interest and the risks for data subjects.[299]

Based on this analysis, various exceptions under Art. 9 (2) GDPR for the processing of sensitive data in the context of the use of AI systems for decisions by public authorities are available. And, indeed, since the exemplary applications are based on such sensitive data, Art. 9 GDPR becomes important and has the potential to prevent discrimination. As seen, Art. 9 GDPR restricts the use of sensitive data but allows exceptions under strict legal conditions. Especially in conjunction with Art. 22 GDPR, which, according to para. 4 prohibits automated decisions based on sensitive data

[296] Ludmila Georgieva and Christopher Kuner, 'Art. 9 Processing of special categories of personal data' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 375.
[297] Marion Albers and Raoul-Darius Veit, 'Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten' in Heinrich Amadeus Wolff, Stefan Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024) para. 57.
[298] Ibid, para. 65.
[299] Ludmila Georgieva and Christopher Kuner, 'Art. 9 Processing of special categories of personal data' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 379.

unless Art. 9 (2) (a) or (g) GDPR is applying, the risk of biased results is minimized. This framework of provisions compels public authorities to deploy AI systems that prevent discriminatory outcomes. A critical point to note, however, is that in the area of big data, there is often the potential to draw inferences from data to sensitive data, thus blurring the line between 'normal' and sensitive data to a certain extent.[300] Yet this is only of secondary importance for the present examination, as the AI sample systems here presumably use sensitive data proactively as the basis for their decisions.

4. Data subject rights under the GDPR

Apart from the described principles, with which all controllers or other responsible parties of AI applications must comply, the GDPR also grants rights to data subjects as set out in Art. 12-23 GDPR, to ensure a comprehensive protection. The data subject has, inter alia, the right to information (Art. 13, 14), the right of access (Art. 15), the right to erasure (Art. 17), the right to data portability (Art. 20 GDPR) and the right not to be subject to automated decision-making (Art. 22). For the present context, a look will first be taken at the information rights before the focus is placed on the right to not be subject to automated decision-making under Art. 22 GDPR.

4.1 Right to Information under the GDPR

The right to information under the GDPR constitutes a cornerstone that is derived from and linked to the fundamental value of transparency.[301] Art. 12 GDPR initially specifies technical and procedural aspects concerning the providing of information.[302] For

[300] Michael Matejek and Steffen Mäusezahl, 'Gewöhnliche vs. sensible personenbezogene Daten: Abgrenzung und Verarbeitungsrahmen von Daten gem. Art. 9 DS-GVO' (2019) 12 ZD 551, 552.
[301] Helena U. Vrabec, *Data Subject Rights under the GDPR* (Oxford 2021) p. 64f.
[302] Radim Polčák, 'Art. 12 Transparent information, communication and modalities for the exercise of the rights of the data subject' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 406.

example Art. 12 (1) stipulates that the information must be provided in a concise, transparent, intelligible and easily accessible form, using clear and plain language and under Art. 12 (2) the controller shall facilitate the exercise of the data subject's rights. The provision thus contains general requirements for a transparent information policy in order to facilitate and simplify the exercise of the rights of data subjects.[303] The fundamental prerequisite for a potential exercise of the rights of the data subject is first of all the knowledge of the processing of personal data, so that Art. 13, Art. 14 and Art. 15 GDPR, which set out the right and access to information, are of key importance.[304]

4.1.1 Information obligations under Art. 13 and Art. 14 GDPR

Art. 13 and 14 GDPR concern the obligation to provide information, but in different situations. Art. 13 GDPR refers to processing activities where personal data is collected from the data subject and Art. 14 GDPR to processing activities where personal data has not been obtained from the data subject directly. Collection means the intentional acquisition of data, and while the data must be collected directly from the data subject for Art. 13 GDPR to apply, some argue this does not preclude it from being collected without the data subject's knowledge.[305] Whereas others require knowledge or cooperation of the data subject for direct collection in the meaning of Art. 13 GDPR.[306] In the area of automated decision-making by AI systems, both forms are initially conceivable with the result that a more specific consideration of one of the two

[303] Lorenz Franck, 'Art. 12 Transparente Information, Kommunikation und Modalitäten für die Ausübung der Rechte der betroffenen Person' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 1.
[304] Rainer Knyrim, 'Art. 13 Informationspflicht bei Erhebung von personenbezogenen Daten bei der betroffenen Person' in Eugen Ehmann and Martin Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024) para. 1.
[305] Alexander Dix, 'Art. 13 Information to be provided where personal data are collected from the data subject' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 4f.
[306] Lorenz Franck, 'Art. 13 Informationspflicht bei Erhebung von personenbezogenen Daten bei der betroffenen Person' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 4 with additional reference.

provisions is not essential. Both provisions set out in para. 1 a catalog of information that must be made available, such as the name and contact details of the controller (Art. 13 (1) (a) and Art. 14 (1) (a)) and the purpose and legal basis for the data processing (Art. 13 (1) (c) and Art, 14 (1) (c)). This catalog is expanded in the respective para. 2 to additional information necessary to ensure fair and transparent processing. The information contained in para. 1 and the information contained in para. 2 are subject to the same obligation that is equally binding, which is made clear by the word “shall”.[307]

In the context of automated AI decision-making, the controller's obligation to inform about the existence of an automated decision-making process including profiling is of particular relevance, as provided by Art. 13 (2) (f) and Art. 14 (2) (g) GDPR.[308] This requires providing information about the fact that an automated decision under Art. 22 (1) or (4) GDPR will be made in the future and about its underlying mechanism, meaning the data subject must be meaningfully informed about the logic involved and the significance and envisaged consequences of such processing.[309] This raises the question of the extent to which information must be provided and whether the underlying algorithm must be disclosed. The logic involved includes a description of the basic principle underlying the structure of this specific automated decision-making process but not the algorithm used, the code lines or how machine learning or algorithmic decision-making works on a general level.[310] In this context, trade secrets

[307] Gabriela Zanfir-Fortuna, ‘Art. 13 Information to be provided where personal data are collected from the data subject’ in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 428.
[308] Parallel to this is also the provision of Art. 15 (1) (h) GDPR within the scope of the right of access.
[309] Alexander Dix, ‘Art. 13 Information to be provided where personal data are collected from the data subject’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 10.
[310] Gabriela Zanfir-Fortuna, ‘Art. 13 Information to be provided where personal data are collected from the data subject’ in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 430; see also Boris P. Paal and Moritz Hennemann, ‘Art. 13 Informationspflicht bei Erhebung von personenbezogenen Daten bei der

can be involved and argue against a publication of the algorithm.[311] The format, the structure and the sequence of basic mechanisms of an automated decision-making process need to be described in order to ensure possible exercising of the rights according to Art. 22 GDPR.[312] This leads to the assumption that the algorithm does not have to be disclosed, despite the question to what extent this would help those concerned. Therefore, the above information may already be sufficient and even lead to a better understanding for the data subject than the provision of the algorithm.

4.1.2 A comprehensive right to information about automated decisions - the right of access under Art. 15 (1) GDPR and Art. 22 (3) GDPR

However, the following considerations would result in a limited degree of effectiveness for these information obligations. The obligations of Art. 13 (2) (f) and Art. 14 (2) (g) GDPR only apply to cases in which an automated decision within the meaning of Art. 22 GDPR is present, i.e. decisions without any human intervention[313], also meaning this obligation does not apply if the systems are only used to support human decision-makers.[314] These transparency requirements are not reduced if a human actor intervenes in an AI scenario by simply adopting the outcome of the decision-making process without adequate review, rather the obligation is reduced if a human assesses the outcome given by the AI before taking the decision.[315] Furthermore, both Art. 13 (2) (f) and Art. 14 (2) (g) GDPR only provide the provision of information and

betroffenen Person' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 31b.

[311] Carsten Orwat, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020), p. 66; see also Andrew D. Selbst, 'The Intuitive Appeal of Explainable Machines' (2018) 87 Fordham L. Rev. 1085, 1092ff. with additional references.

[312] Alexander Dix, 'Art. 13 Information to be provided where personal data are collected from the data subject' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 10.

[313] Mario Martini, 'Algorithmen als Herausforderung für die Rechtsordnung' (2017) 21 JZ 1017, 1020.

[314] Thomas Wischmeyer, 'Regulierung intelligenter Systeme' (2018) 143 AöR 1, 50.

[315] Alexander Dix, 'Art. 13 Information to be provided where personal data are collected from the data subject' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 10.

explanation prior to the decision, the notification takes place before a decision is made, by the point when the data is collected, which means this does not constitute a right to explanation.[316] Specific information regarding a particular decision can - logically - only be provided and displayed after the decision has been made, meaning only general information about how the system functions has to be provided. Although Art. 22 (3) GDPR and Recital 71 provide the data subject to obtain an explanation of the decision reached after such assessment, in order to be able to challenge the assessment[317], this right is not mentioned in the respective paragraphs of Art. 13 or Art. 14 GDPR and thus could be regarded as non-binding.[318] Certain information could then only be provided voluntarily after the decision was made as part of possible safeguards.[319] In this regard, it should be noted that Art. 15 (1) (h) GDPR, which provides the right of access in the case of automated decisions with the same wording as the respective paragraphs in Art. 13 and Art. 14 GDPR, has no deadline, meaning it appears possible to assert the right even after the decision has been made. But in view of the similar system to Art. 13 and 14 GDPR, and the future oriented terminology of 'envisaged consequences' used in Art. 15 (1) (h) GDPR the assumption that this right covers the reasons and causes for a decision that has already been taken is criticized.[320] Based on these considerations, it would have to be assumed that the GDPR does not provide a possibility to obtain information about the decision ex post - i.e. a right to explanation.

[316] Sandra Wachter, Brent Mittelstadt, et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 Int. Data Priv. Law 76, 82
[317] See Rec. 71 Sentence 4 GDPR.
[318] Sandra Wachter, Brent Mittelstadt, et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 Int. Data Priv. Law 76, 82.
[319] Sandra Wachter, Brent Mittelstadt, et al., 'Counterfactual Explanations without Opening the Black Box: Automated Decision and the GDPR' (2018) 31 Harv. J. L. & Tech. 841, 862.
[320] Sandra Wachter, Brent Mittelstadt, et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 Int. Data Priv. Law 76, 83.

However, in view of the overarching protection objective of the GDPR, this is not very convincing. The Guidelines on Automated individual decision-making of the Data Protection Working Party made clear that there is a need for the data subject to know on what basis and with what weighting the decision was made in order to exercise their own rights, meaning the suitable safeguards in Recital 71 include these information.[321] The purpose of Recital 71 in connection with Art. 22 (3) GDPR, which provides for the suitable safeguards, as well as the information obligations under Art. 13 and 14 GDPR and the right to access under Art. 15 GDPR, is to ensure that the data subject is informed to have the possibility to contest such a decision. The content, scope and timing of the information that needs to be provided is determined with this objective in focus.[322] The information must be meaningful for the data subject, i.e. the information about an automated decision must at least be sufficient to enable the data subject to recognize whether it has an actionable discrimination claim.[323] Accordingly, the right to information regarding the decision that has already been made and the underlying reasons must be seen at least in Art. 15 (1) (g) GDPR.[324] Furthermore, although Recital 71 is not binding, its importance is not minor, as it is to be used for interpretation and thus supports the imprecise text of the GDPR in the direction of a more in-depth right of explanation. Considering the overall objective of the GDPR and the fundamental principle of transparency, the data subject can receive information even after the

[321] Art 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' 17/EN WP251rev.01, p. 27.
[322] Margot E. Kaminski, 'The Right to Explanation, Explained' (2019) 34 Berkley Tech. L. J. 189, 211f.
[323] Andrew D Selbst and Julia Powles, 'Meaningful information and the right to explanation' (2017) 7 Int. Data Priv. Law 233, 236.
[324] See Alexander Dix, 'Art. 15 Right of access by the data subject' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 19; see also Gianclaudio Malgieri and Giovanni Comande, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7 Int. Data Priv. Law 243, 255f.

decision has been made and thus also obtains the factors on which the decision is based.

As a conclusion, it can be stated that there is a comprehensive ex post right to explanation in relation to automated decisions by AI systems, namely in Art. 15 (1) (g) GDPR and Art. 22 (3), as the necessary safeguards imply an indirect obligation to explain the decision-making process. These rights ensure individuals are informed about how AI systems make decisions and help to uncover discriminatory patterns and highlighting potential biases in the process if the underlying and especially the decisive factors for the decision are presented in a comprehensible manner. Irrespective of this legal basis for a comprehensive information obligation, however, the technical feasibility must also be taken into account and is very questionable based on the explanations regarding the transparency of AI systems.[325] Yet, the discussion of this exceeds the scope of this paper, which only aims at providing a legal analysis.

4.1.3 The state's obligation to provide reasons

In addition to these options for obtaining information provided by the GDPR, the obligation of public authorities to state reasons is also relevant in the present context. Administrative law obliges authorities to justify their decisions[326], which means they must explain how AI has influenced the outcome. Thus, it can be assumed this obligation further ensures that individuals can understand and challenge AI-based decisions to counteract discrimination.[327]

[325] From such a perspective it can be argued against an advantage for the affected person due to the lack of possibilities to provide a technically correct explanation of the specific decision made by an AI system, see e.g. Thomas Wischmeyer, 'Regulierung intelligenter Systeme' (2018) 143 AöR 1, 53.
[326] In Germany laid down in § 39 Verwaltungsverfahrensgesetz.
[327] Further explanations on this legal possibility and obligation to provide information are omitted here as the analysis focuses primarily on the GDPR.

4.2 Automated decision making, Art 22 GDPR

Art. 22 GDPR establishes the right not to be subject to a decision based solely on automated processing - including profiling[328] - which produces legal or similarly significant effects and is therefore of particular relevance in relation to AI systems in the present context. This provision is an expression of the objective that processing of personal data should be designed to serve mankind[329], in that mankind is not subject to fully automated decisions. Recital 71 also refers directly to the potential risks of discrimination by requiring personal data to be secured in a manner that takes into account the potential risks to the interests and rights of the data subject and to prevent, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or belief, trade union membership, genetic or health status or sexual orientation, or processing that results in measures having such an effect.[330] The provision is therefore based on the high risk the legislator recognized in such automated decisions.[331] Art. 22 GDPR is not equivalent or comparable to the other provisions of the GDPR, as it is not based on the processing, but rather focuses on the consequences and outcome of an automated decision.[332]

Art. 22 (1) stipulates a general prohibition of automated decision-making for individual cases while Art. 22 (2) and - for sensitive data - Art. 22 (4) provide exceptions to this

[328] Profiling is defined in Art. 4 (4) GDPR as follows: 'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.
[329] See Rec. 4 Sentence 1 GDPR.
[330] See Rec. 71 Sentence 6 GDPR.
[331] Sebastian Schulz, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 1.
[332] Tristan Radtke, 'Das Verhältnis von KI-VO und Art. 22 DS-GVO unter besonderer Berücksichtigung der Schutzzwecke' (2024) 8 RDi 353, para. 4; see also Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 1.

rule. Art. 22 (3) requires suitable measures to safeguard the rights and freedoms as well as the legitimate interests if one of these exceptions applies. Although Art. 22 (1) GDPR explicitly refers to a 'right' and is also in the corresponding section of the GDPR, it does not establish a genuine right. Rather it is to be understood as a general prohibition without requiring any specific action from the data subjects.[333] It therefore constitutes a prohibition which is independent of individual enforcement.[334] This results from its independent role alongside Art. 21 GDPR; if it were merely a right of the data subject, the provision would basically be obsolete alongside Art. 21 GDPR, which does not seem to be intended by the legislator.[335] In addition, a right to object would contradict the possibility of consent set out in para. 2 lit. c, as the data subject cannot object and consent to the same processing at the same time.[336] Finally, the Guidelines of the Working Party on Automated individual decision-making point out that even though the word 'right' is used, this is to be understood as a general prohibition.[337]

4.2.1 Scope of application – classification of the use cases

To determine the extent to which the examples examined here fall within the scope of Art. 22 GDPR, it is necessary to look at the scope of application. Art 22 (1) has three fundamental requirements for application: a decision must have been taken; it must be

[333] Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 52; see also Olivia Tambou, 'Art. 22 Automated individual decision-making, including profiling' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 7.
[334] Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 29b; see also Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 59.
[335] Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 29b.
[336] Olivia Tambou, 'Art. 22 Automated individual decision-making, including profiling' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 8.
[337] Art 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' 17/EN WP251rev.01, p. 19.

based exclusively on automated processes; and it must cause legal or similarly significant effects.[338]

The first requirement is accordingly a decision. This must be a decisive act with a certain conclusive effect, which is legally assigned to an individual person, i.e. relate to an individual case.[339] But the term decision is to be seen quite broadly, as measures are also included[340], so that it essentially means a certain attitude or stance is taken towards a person and this attitude/stance has a certain binding effect in the sense that it is likely to be followed.[341] Whether or not such a decision has to involve a certain degree of complexity is irrelevant here, as the use of AI systems fulfills this requirement either way.[342]

In the area of law enforcement, as mentioned, the Law Enforcement Directive is partly relevant. However, Art. 11 (1) LED contains a parallel provision on automated individual decision-making, though this provides a wider range of exceptions and is therefore more liberal. Irrespective of this, the latter one too requires a decision. Particularly in the case of the COMPAS system, which predicts a defendant's likelihood of recidivism, this requirement could be missing, as only a prediction is made at first, but no decision is taken and the mere calculation of a probability does in principle not

[338] See Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 43; see also Olivia Tambou, 'Art. 22 Automated individual decision-making, including profiling' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 10.
[339] Kai von Lewinski, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Heinrich Amadeus Wolff, Stefan Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024) para. 14f.
[340] See Rec. 71 Sentence 1 GDPR; see also Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 45.
[341] Isak Mendoza and Lee A. Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou, Philippe Jougleux, et al. (eds), *EU Internet Law: Regulation and Enforcement (*Springer 2017) p. 87.
[342] On this question, see e.g. Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 15b.

constitute a decision in this sense.[343] Although an assessment of personal aspects relating to a natural person with a predictive character is carried out here, which means that profiling may be assumed, this is only covered by Art. 22 (1) GDPR if it leads to an individual decision.[344] The results of the system are used as a basis for the judgment, the sentence and an application for early release, but this happens in a second step, meaning the system itself does not make a decision.[345] It always depends on the individual case, but considering the broad interpretation of the condition 'decision' by the ECJ in the SCHUFA Holding case from December 2023 on Art. 22 GDPR, this condition can be assumed here if a criminal decision “draws strongly”[346] on this basis.[347] In the context of predictive policing systems, however, a decision within the meaning of Art. 22 (1) GDPR can rather not be assumed, as in most cases no direct decision is made as the initial aim is to identify potential criminal offenses, i.e. this does not fall within the scope of Art. 22 (1) GDPR. When AI systems are deployed at the EU borders, this may look a little different; although systems are likewise used in situations in which no direct individual decision is made, it is conceivable that automated questioning and biometric verification systems could be used as the decisive basis for a decision to refuse entry, meaning the decision “draws strongly” on the solely automated processing.

The AMS system is a scoring system that uses a database to make predictions about individuals' chances of integration in the labor market and thus again can be qualified

[343] Sebastian Schulz, ‘Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling’ in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 17.
[344] See Helena U. Vrabec, *Data Subject Rights under the GDPR* (Oxford 2021) p. 192; see also Mario Martini, ‘Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling’ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 23.
[345] See Julia Angwin, Jeff Larson, et al., ‘Machine Bias’ (*propublica* 23 May 2016) <propublica.org> accessed 3 Dec 2024.
[346] Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 73.
[347] Ibid, para. 44f.

as profiling in the sense of the GDPR. Here again, reference can be made to the ECJ's SCHUFA ruling.[348] Although SCHUFA only determined a certain score and a decision was taken by a third party later, the term was found to be broad enough to encompass the result of the calculation of a person's creditworthiness in the form of a probability value about that person's ability to meet payment obligations in the future.[349] Accordingly, with regard to the AMS system, it can also be assumed that this requirement is met when job proposals and measures are drawn strongly on the basis of the score determined. Finally, the application in social welfare and, in this respect, the detection of fraud as well as the application for study place allocation can clearly be subsumed under this, as decisions are made with regard to individuals, such as the granting of social welfare benefits or the acceptance of a study place.

The second requirement, that this decision must be based exclusively on automated processes is slightly more difficult to classify. Firstly, this includes all decisions that are made without any human influence.[350] While every automated decision is to some extent subject to minimal human influence, as they have usually designed the technology, such as the underlying algorithm, this influence in the development stage can probably not be taken into account, as the scope of Art. 22 (1) GDPR would otherwise be very limited.[351] Thus, it should rather be understood in the sense of a solely automated decision if a human has no genuine influence on the outcome of the decision-making process.[352] Similarly, it is not enough for a person to merely look at

[348] Ibid.
[349] Ibid, para. 46; see also Rec. 71 Sentence 1 GDPR.
[350] Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 16.
[351] Olivia Tambou, 'Art. 22 Automated individual decision-making, including profiling' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 14.
[352] Lee A. Bygrave, 'Art. 22 Automated individual decision-making, including profiling' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 532.

the result without any substantive examination or in other words, it is not enough to simply convert a certain evaluation of the computer into a decision.[353] The Guidelines on Automated decision-making state the controller must ensure that any review of the decision is meaningful and not just a token gesture, and further that it should be carried out by a person who has the authority and competence to change the decision.[354] However, there is in principle no such decision based solely on automated processing if the system is only used to support the decision-making process.[355] Although the ECJ extended the scope of application to a certain extent through the SCHUFA ruling, as it interprets the requirements broadly and states that it is enough if the decision "draws strongly" on the value or the score issued by automated processing, meaning it could be assumed decision support systems are also covered as long as this requirement is met.[356] The court does not state when there is such drawing, but it can be concluded from the reasoning that Art. 22 (1) applies if a pattern can be recognized, such as no social benefits are paid in the case of a bad score as the court stated "an insufficient probability value leads, in almost all cases, to the refusal of that bank to grant the loan".[357] However, it cannot be concluded directly from this that support systems are also covered, since, firstly, it is not sufficiently clear at what point a person draws strongly on the results of automated processing - given the wording 'solely', from a

[353] Kai von Lewinski, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Heinrich Amadeus Wolff, Stefan Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024) para. 24ff.
[354] Art 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' 17/EN WP251rev.01, p. 21.
[355] Lee A. Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making', in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford, 2019; online edn) p. 253; see also Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 59f.; Isak Mendoza and Lee A. Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou, Philippe Jougleux, et al. (eds), *EU Internet Law: Regulation and Enforcement* (Springer 2017) p. 87.
[356] Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 73, see Jan Horstmann, 'CJEU: The Rating of a Natural Person's Creditworthiness by a Credit Rating Agency Constitutes Profiling and Can Be an Automated Decision under Article 22 GDPR' (2024) 1 Eur. Data Prot. Law Rev. 117, 120.
[357] Case C-634/21 OQ v Land Hessen [2023] ECLI:EU:C:2023:957, para. 48.

technical legal point of view this is to be interpreted rather narrowly - and, secondly, it cannot be determined to what extent a person has actually relied on the information or has included their own assessments. As seen, the threshold at which human intervention rules out a solely automated process is not easy to determine. When using AI systems for decision-making in the cases in question, this requirement poses major problems for the applicability of Art. 22 GDPR. Although it depends on the individual case and application, it can be assumed a decision made by the systems in focus here is reviewed by a human at the end or that the systems merely act as a decision-making support where the outcome is only the basis of a final decision made by a natural person. This would rule out the application of Art. 22 GDPR, thereby resulting in a situation where the prohibition of Art. 22 (1) GDPR does not apply, even though there is a significant risk that the evaluations and outcomes of the AI are blindly followed without carrying out an own evaluation.[358] Something different can only be considered if, e.g. in the area of the allocation of study places certain applicants are sorted out directly by an automated process and thus do not come under human consideration at all.[359] However, it is different if the AI application only carries out a ranking and on this basis a human decides which applicants are possible candidates.[360] It is therefore highly dependent on the individual case whether such a solely automated processing exists, but in the cases presented here it is most likely that no such solely automated processing occurs, since the systems mentioned serve more to prepare public decisions than to make final decisions themselves. This means that the protection of Art. 22 GDPR does not apply, so the most important provision of the GDPR offers very

[358] See e.g. Lee A. Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making' in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford, 2019; online edn) p. 253.
[359] Jakob Hüger, *Künstliche Intelligenz und Diskriminierung* (Nomos 2023) p. 362.
[360] Sebastian Schulz, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 12.

limited protection, considering that many decisions or recommendations are simply adopted without further examination by a human being and are nevertheless not subject to Art. 22 GDPR. Irrespective of this, it can be assumed that the last requirement, that the decision produces legal effects or similarly significantly affects, will always be present in decisions of public authorities as examined here.[361]

Based on this analysis, a very limited scope of application can be determined in the present context. Looking to the future, however, it can be assumed that the use of AI systems will also increase in areas of administration and expand towards a solely automated processing. Additionally, this can be assumed by considering that humans may not have access to all the information that the AI has included and may not be able to review how that information has been used, making it overly burdensome to conduct an effective review.[362] Therefore, the protection mechanisms of Art. 22 GDPR shall be outlined briefly.

4.2.2 Protection mechanism of Art. 22 GDPR regarding discrimination

Firstly, Art. 22 (2) GDPR states three cases in which an automated individual decision that has a legal consequence is nevertheless permitted. Although they are exhaustive[363], these exceptions are quite broad[364], which further limits the scope of protection of the provision. The exceptions laid down in Art. 22 (2) (a)-(c) GDPR are that either it is necessary for the conclusion or performance of a contract; it is based

[361] Although these only include those that establish, change or revoke a legal position and have a negative effect on the person concerned, see Sebastian Schulz, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Peter Gola and Dirk Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022) para. 21.
[362] Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 60.
[363] Mario Martini, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 30.
[364] Kai von Lewinski, 'Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Heinrich Amadeus Wolff, Stefan Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024) para. 42.

on a legal authorization with parallel safeguards in favor of the data subject; or the explicit consent of the data subject was given. Within the scope of derogation under Art. 2 (2) (b) GDPR (authorized by Union or Member State law), the Member States have a relatively wide margin of discretion, especially since the required appropriate measures for the protection of data subjects are formulated in relatively abstract terms.[365] The suitable measures to safeguard the data subject's rights, freedoms and legitimate interests are not aligned with those of Art. 22 (3) GDPR, as this actively refers only to the exceptions of Art. 22 (2) (a) and (c).[366] Although it can be assumed that Art. 22 (3) GDPR sets the objective and a lower level of protection would be unlawful, the term is rather broad to understand so Member States have a relatively large degree of autonomy and a wide range of protective measures to take.[367]

Regarding the other two exceptions, i.e. contract and consent, Art. 22 (3) GDPR provides a qualification.[368] Similar to the aforementioned Art. 22 (2) (b) GDPR, this provides for suitable measures to safeguard the data subject's rights, freedoms and legitimate interests, but additionally provides for minimum criteria such as “at least” the right to human intervention on the part of the controller, to express his or her point of view and to contest the decision. Accordingly, this list is not exhaustive but only represents the minimum set of rights a data subject is supposed to have.[369] On the

[365] Lee A. Bygrave, ‘Art. 22 Automated individual decision-making, including profiling’ in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 537.
[366] Mario Martini, ‘Art. 22 Automatisierte Entscheidungen im Einzelfall einschließlich Profiling’ in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 35a.
[367] Olivia Tambou, ‘Art. 22 Automated individual decision-making, including profiling’ in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 29.
[368] Lee A. Bygrave, ‘Art. 22 Automated individual decision-making, including profiling’ in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 538.
[369] Lee A. Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision-Making' in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford, 2019; online edn) p. 254.

question of whether a right to an explanation after decision-making arises from this provision (also in connection with Recital 71), reference is made to the discussion above.[370] In any case, it is positive to emphasize that Art. 22 (3) GDPR - at least in the cases of applicability of para. 1 and the corresponding permission in para. 2 - provides a framework to counteract discrimination by AI systems, through transparency and information rights, human oversight, and the right to challenge the decisions. However, it is questionable how strong these possibilities, which are at least laid down in law, are, as they rely heavily on practical implementation, technical capabilities, and systemic enforcement.

Finally, Art. 22 (4) GDPR stipulates that automated decisions shall not be made based on sensitive categories of personal data, even in the exceptional cases of para. 2. This qualified prohibition[371] again recognizes two exceptions. Namely that the data subject has either given explicit consent to the processing for one or more specified purposes according to Art. 9 (2) (a) GDPR or that the processing is carried out for reasons of important public interest based on Union or Member State law according to Art. 9 (2) (g) GDPR. In addition, para. 4 also requires suitable measures to safeguard the data subject's rights, freedoms and legitimate interests. Although the exact scope is not specified, it is likely to be interpreted in the same way as the measures referred to in Article 22 (3) GDPR.[372] This general prohibition in itself helps to prevent discrimination by AI systems in any case, since no sensitive data is allowed to be used, on which discrimination is often based. However, there are also limitations here, on the one hand obviously the exceptions of the provision as such. But on the other hand, specifically

[370] See VI.4.1.2.
[371] Lee A. Bygrave, 'Art. 22 Automated individual decision-making, including profiling' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 539.
[372] Ibid.

in relation to AI systems the fact that sensitive data can often be derived from normal data as well as the problem of proxy discrimination, meaning AI systems can indirectly infer sensitive data from non-sensitive attributes.[373] Furthermore, the scope of sensitive data under the GDPR by no means covers all the characteristics that often give rise to discrimination.

Regarding Art. 22 GDPR, the following can be stated in conclusion: At first glance, Art. 22 GDPR does offer an in-depth protection against discrimination by AI systems through public authorities by prohibiting ADM in principle. On closer examination, however, this protection is very fragmentary, as these applications usually do not even fall within the scope of Art. 22 (1) GDPR. In addition, the suitable measures required by both Art. 22 (2) (b) and Art. 22 (3) GDPR are not entirely easy to determine or even impossible to implement when using AI applications. Finally, and probably the most important fact is that the declaration of a final human review is easily done making it no longer a solely automated process, although in reality this does not happen at all. Either for technical reasons it is not possible to review the AI decision accurately or it does not happen for the simple reason that humans tend to adopt these results without further consideration. It should still be noted, however, that Art. 22 GDPR provides actual protection for systems that - especially in the future - will fall under this prohibition and, assuming the right to explanation, it contributes at recognizing and combating discrimination.

[373] See Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 62.

4.3 Further obligations when processing personal data with relevance to discrimination

Finally, the further obligations when processing personal data for the controller and processor are analyzed shortly with a view to their contribution against discrimination.

4.3.1 General compliance obligation and Data protection by design and by default, Art. 24 and 25 GDPR

Art. 24 GDPR firstly sets out the general obligation of the controller to comply with the provisions of the GDPR. The controller must take appropriate and effective measures and be able to demonstrate the compliance of processing activities[374], which includes addressing potential discrimination in AI systems.[375] The determination of which organizational and technical measures are necessary is not specified or linked to a minimum threshold, para. 1 merely requires that "the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons" must be considered.[376] The scope of the measures thus depends on the potential risks for the data subject, i.e. a risk-based approach is the underlying principle.[377] For AI systems that are supposed to make or support decisions, this means measures are necessary with regard to training data, especially in terms of completeness, as well as measures to detect and remove bias in the algorithms to prevent discrimination by regularly monitoring AI performance.[378]

[374] See Rec. 74 Sentence 2 GDPR.
[375] See Rec. 75 GDPR.
[376] Stephan Schmidt and Stefan Brink, 'Art. 24 Verantwortung des für die Verarbeitung Verantwortlichen' in Heinrich Amadeus Wolff, Stefan Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024) para. 16.
[377] Jos Dumortier and Pieter Gryffroy, 'Art. 24 Responsibility of the controller' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 23.
[378] See Giovanni Sartor and Francesca Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020) p. 67.

Art. 25 (1) and (2) GDPR contain the concepts of data protection by design and data protection by default. The controller is thereby obliged to integrate technical and organizational measures designed to effectively implement the data protection principles and to integrate necessary safeguards, during as well as prior to processing, when creating the system.[379] Art. 25 (1) GDPR and also Recital 78 remain quite vague with regard to specific measures, but this is due to the fact that these measures concern the entire operating cycle and numerous risks; Art. 25 (2) GDPR, on the other hand, provides four more specific elements to be considered to limit the processing of personal data by default.[380] Thus, AI systems must be designed to minimize bias, with representative data sets to reduce the risk of discriminatory decisions. In addition, the default settings must ensure that only data required for the specific purpose is processed to prevent the misuse of any personal data.

4.3.2 Data Protection Impact Assessment, Art. 35 GDPR

Finally, Art. 35 GDPR provides for a Data Protection Impact Assessment if processing is likely to result in a high risk to the rights and freedoms of natural persons and especially when using new technologies. This serves as an early warning mechanism, to detect possible legal violations through the processing in order to take appropriate measures.[381] According to Art. 35 (3) (a) GDPR, a DPIA is necessary if a systematic and comprehensive evaluation of personal aspects relating to natural persons is based on automated processing, including profiling, and on which decisions are based that produce legal effects or similarly significantly affect a natural person. In addition, a

[379] Lee A. Bygrave, 'Art. 25 Data protection by design and by default' in Christopher Kuner, Lee A Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020) p. 576.
[380] Marco Almada, Juliano Maranhao, et al., 'Art. 25 Data protection by design and by default' in Indra Spiecker gen. Döhmann, Vagelis Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023) para. 16f.
[381] Mario Martini, 'Art. 35 Datenschutz-Folgenabschätzung' in Boris P. Paal and Daniel A. Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021) para. 6.

DPIA is always required if special categories of data within the meaning of Art. 9 (1) GDPR or personal data relating to criminal convictions and offenses within the meaning of Art. 10 GDPR are processed on a large scale. Firstly, it can be assumed that since the wording only refers to “on which decisions are based” and not to a “decision based solely on automated processing” as in Art. 22 (1) GDPR, the scope here is wider.[382] Thereby, it includes situations in which a human is involved in the decision and can therefore also include the systems discussed here. Secondly, it can be assumed that usually - if not always - sensitive data within the meaning of Art. 9 GDPR is processed, so that a DPIA would always be mandatory here as well. Thus, it ensures that AI systems are designed and operated in compliance with GDPR principles, thereby reducing the risk of discrimination.

5. Interim findings – GDPR as a safeguard against discrimination through AI systems

Overall, in terms of the opportunities provided by the GDPR to minimize discrimination by AI systems, the following can be stated: First of all, the most important requirement for effective AI is the underlying dataset, i.e. the more data the better the training, what generally contradicts the GDPR. In the field of AI for decision-making or support by public authorities discussed here, personal data is always involved, as decisions are made in relation to specific individuals, which means that the provisions of the GDPR must be observed. This means the far-reaching principles and requirements must be complied with, with transparency and information obligations in particular encountering the problem of the opacity of AI. At the same time, Art. 22 GDPR, as the main provision concerning automated AI decisions, provides rather incomplete protection as the

[382] See Art 29 Data Protection Working Party, ‘Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679’ 17/EN WP251rev.01, p. 29.

scope of application is already limited and further far-reaching exceptions apply. In other words, it's double-edged: on the one hand, AI systems are almost over-regulated by the GDPR, but on the other hand, such application fields are not covered by some of the regulations. Personal data and, in particular sensitive data within the meaning of Art. 9 GDPR are necessary to identify discrimination, since, e.g. a certain ethnicity must be known in order to test whether the AI system has discriminated on this basis.[383] Therefore, it is questionable to what extent the GDPR prevents an effective protection against discrimination. Although the GDPR provides legal possibilities, even apart from Art. 22 GDPR, which can prevent discrimination, these are mostly difficult to implement from a technical point of view. However, the SCHUFA ruling provides a direction for the first time, namely that Art. 22 GDPR must always be interpreted with regard to the rights of the data subjects by assuming a rather broad scope of application. Although it is not yet clear to what extent support systems, such as those in question here, are actually covered, this direction is now quite conceivable.[384] This approach of the ECJ can to a certain extent also be applied to the other issues discussed here, which are still controversial, with the result that the question of a right to an explanation as well as the question of suitable measures is more likely to be answered in the interests of the data subjects. However, it should also be noted that non-discrimination law of course applies independently of the GDPR to any case of discrimination and provides subsequent protection mechanisms but are not subject to this analysis.[385] It therefore remains to be examined to what extend the AI Act

[383] Marvin van Bekkum and Frederik Zuiderveen Borgesius, 'Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?' (2022) 48 Comput. L. & Secur. Rev. 105770, p. 3.
[384] Similar opinion, see Jan Horstmann, 'CJEU: The Rating of a Natural Person's Creditworthiness by a Credit Rating Agency Constitutes Profiling and Can Be an Automated Decision under Article 22 GDPR' (2024) 1 Eur. Data Prot. Law Rev. 117, 121ff.
[385] In addition, the general framework of fundamental rights, the EU Charter of Fundamental Rights and the European Convention on Human Rights apply to any use of AI.

counteracts discrimination by AI systems, as, in comparison to the GDPR, it was tailored precisely to such systems.

VII. Legal protection under the AI Act

As already pointed out in the introductory part of the AI Act, it pursues two main objectives: the fostering of innovation on the one hand and the protection of individuals and fundamental rights on the other hand.[386] To achieve these objectives, the AI Act relies on a risk-based approach with four levels, which are unacceptable, high, limited or minimal risk. The questions are therefore: How do the AI systems in question here in the context of decision-making need to be classified under the AI Act? To what extent does the law provide measures for protection against discrimination and for the protection of fundamental rights?

1. Classification of AI systems for decision-making under the AI Act

The fundamental prerequisite for the application of the AI Act is, of course, the existence of an AI system within the meaning of Art. 3 (1) AI Act. Differentiating between the existence of a 'mere' algorithm or set of algorithms from the existence of an AI is quite complex. However, the broad definition of AI systems in the AI Act addresses this problem by covering all machine-supported, autonomous operating systems that have the ability to infer and generate outputs such as predictions, content, recommendations or decisions, Art. 3 (1) AI Act. This means in principle even complex data processing can be classified as an AI system as long as it produces one of the stated outputs. Thus, the systems in consideration here would fall within the scope of the AI Act, as an AI system within the meaning of the AI Act covers a wide range of

[386] See V.2.

systems. Regarding the AMS system and the Parcousup system, the applicability in its current form may be questioned as it could be classified as a pure algorithm. The AMS tool is based purely on an algorithm. So, this system is rather not to be classified as AI in the sense of Art. 3 (1) AI Act.[387] Parcoursup uses algorithms to process and rank applications and does so in a machine-supported and automatic way, albeit according to pre-programmed rules, which is, according to Recital 12 not covered.[388] However, it is very conceivable that in the future such systems with AI features will be used in this area. Based on this finding, it is therefore necessary to determine in which risk category the systems are to be classified.

1.1 Law Enforcement Systems under the AI Act

The COMPAS system, that predicts a defendant's likelihood of recidivism, could constitute a prohibited AI technique under Art. 5 (1) (d) AI Act. In the proposed Act, this category of systems was not listed as a prohibited system under Art. 5, which means on this basis such systems were classified as high-risk systems in any case.[389] However in the final text, AI systems for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offense, based solely on the profiling of a natural person or on assessing their personality traits and characteristics, are prohibited. For the definition of profiling, Art. 3 (52), the AI Act refers to Art. 4 No. 4 GDPR, meaning it includes the automated processing of personal data to evaluate certain personal aspects relating to a natural person. Even if profiling is permitted under the GDPR in specific cases, an assessment of the likelihood that a

[387] See Hannah Ruschemeier, 'Hochrisiko-KI-Systeme gemäß Artikel 6 Absatz II' in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 51.
[388] See Rec. 12 Sentence 2 AI Act.
[389] See Gijs van Dijck, 'Predicting Recidivism Risk Meets AI Act' (2022) 28 Eur. J. Crim. Policy Res. 407, 410ff.

person will commit a crime is still prohibited.[390] This prohibition is intended to ensure the presumption of innocence by only allowing people to be judged based on objective grounds.[391] Yet, this does not include all predictive policing systems, only those that are based solely on the profiling of a natural person or the assessment of their personality traits and characteristics. Recital 42 clarifies that it is prohibited to base an AI predicted behavior solely on profiling, personality traits or characteristics "without a reasonable suspicion that the person being involved in a criminal activity based on objective, verifiable facts and without human assessment".[392] Moreover, the provision itself exempts the use of AI systems supporting the human assessment of a person's involvement in a criminal activity that is already based on objective and verifiable facts directly related to a criminal activity. This raises the question of what is to be understood by "without reasonable suspicion". A reasonable suspicion could already be assumed because it concerns the likelihood of reoffending, meaning that there is already a connection to the commission of a criminal offense. But such an interpretation is specifically not in line with the presumption of innocence, which of course also applies again after every criminal offense committed and on which the ban is fundamentally based. The question also arises as to when this occurs without human assessment. Here, reference can partly be made to the explanations on Art. 22 GDPR, as the problem of the unclear threshold is present again, as to when a human intervention of this kind is given.[393] Therefore, whether a system such as COMPAS on the basis of which a decision is made by a human would already meet this requirement. Irrespective of these considerations, however, several factors speak against the assumption that a system such as COMPAS would be subject to Art. 5 (1) (d) AI Act.

[390] Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 43.
[391] See Rec. 42 Sentence 1 AI Act.
[392] See Rec. 42 Sentence 2 AI Act.
[393] See Jessie Levano, 'Predictive Policing in the AI Act: meaningful ban or paper tiger?' (*European Law Blog* 5 July 2024) <europeanlawblog.eu> accessed 3 Dec 2024.

First, the system does not indicate whether a person will commit a specific offense in the future, but rather issues a score that is intended to assess the likelihood of recidivism in general. Second, the prohibited practice, as described by the scholars, relates specifically to systems of predictive policing, i.e. systems to forecast criminal activity.[394] Accordingly, it covers crime prevention tools, i.e. predictive policing systems. The COMPAS system, however, is not a crime prevention tool in the narrower sense because based on the issued probability no criminal offenses are to be avoided through criminal prosecution measures in the preventive sense; instead, it serves as a basis for exercising repressive measures. Third, the list of prohibited practices in Art. 5 (1) AI Act is to be interpreted narrowly due to its far-reaching nature and the fact that it deviates from the actual regulatory objective of the AI Act as a product safety law by prohibiting certain AI uses.[395] In view of this consideration, it could be argued that a system such as COMPAS would not fall under the prohibition of Art. 5 (1) (d) AI Act. However, in view of the wording, initially the prohibition appears to be applicable to systems such as the COMPAS system, as it predicts the probability of a person committing a crime again and does so based on profiling as stated above. The implementation of this prohibition may also have been inspired directly by the discussions about the COMPAS system, meaning the legislator wanted to include it.[396] For comprehensive protection of fundamental rights, it is in any case necessary to include such an AI system in this category. Nevertheless, it always depends on the individual system and again on the extent to which human assessment is carried out.

[394] The prohibition of Art. 5 (1) (d) AI Act is subsumed under "predictive policing", see e.g. Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 42; Jessie Levano, 'Predictive Policing in the AI Act: meaningful ban or paper tiger?' (*European Law Blog* 5 July 2024) <europeanlawblog.eu> accessed 3 Dec 2024.
[395] Christoph Krönke, 'Das europäische KI-Gesetz: Eine Verordnung mit Licht und Schatten' (2024) 8 NVwZ 529, 532.
[396] Christiane Wendehorst, 'Art. 5 Verbotene Praktiken im KI-Bereich' in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 72.

Regarding other predictive policing systems, the following can thus be stated. Prohibited in this context are AI systems that are used for predictive identification, i.e. to anticipate criminal offenses in relation to certain people.[397] In this case, also only if it is based solely on the profiling of a natural person or on assessing their personality traits and characteristics. In contrast, this states that systems which do not relate to a certain person but rather to the probability of criminal offenses in a certain area, such as predictive mapping, are not included.[398] Likewise, it is not prohibited to use an AI system to provide support for the human assessment of a person's involvement in a criminal offense if the possible involvement is already based on objective and verifiable facts that are directly related to a particular criminal activity.[399] It therefore also depends on the specific system of predictive policing, but the scope of the prohibition is questionable since the extent of human influence that leads to the prohibition not being applied is unclear and therefore systems can easily fall outside the scope of the prohibition.[400] In the context of law enforcement, Art. 5 (1) (h) AI Act may also be relevant, which prohibits the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement. While Recital 32 explicitly justifies this with the risk of biased and discriminatory outcomes, there are however quite far-reaching exceptions to this prohibition in Art. 5 (1) (h) AI Act.[401] Moreover, according to the wording, only 'real-time' identifications are covered which means under Art. 3 (42) AI Act the collection of biometric data, comparison and

[397] Jessie Levano, 'Predictive Policing in the AI Act: meaningful ban or paper tiger?' (*European Law Blog* 5 July 2024) <europeanlawblog.eu> accessed 3 Dec 2024.
[398] See Nikos Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act* (Springer 2023) p. 384.
[399] Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 43.
[400] Jessie Levano, 'Predictive Policing in the AI Act: meaningful ban or paper tiger?' (*European Law Blog* 5 July 2024) <europeanlawblog.eu> accessed 3 Dec 2024.
[401] See Rec 32 Sentence 2 AI Act.

identification without significant delay, which excludes post systems entirely and also online live biometric identification, e.g. via video streams.[402]

Consequently, it can be found that such law enforcement systems are not prohibited - except for the COMPAS system, which is difficult to classify, though more likely to be considered a prohibited practice - but can instead be classified as stand-alone high-risk systems under Art. 6 (2) AI Act. Art 6 (2) AI Act classifies all AI systems listed in Annex III as high-risk systems. Such systems are here subject to No. 6 (d) (law enforcement, insofar as their use is permitted under relevant Union or national law).

In the area of law enforcement, the exception concerning national security under Art. 2 (3) AI Act is additionally relevant, as the AI Act would not apply at all if the AI deployment were considered a matter of national security by an EU member state.[403] It is not clear what can be subsumed under the national security exception, meaning that there could be a risk that Member States use it to justify the use of predictive identification techniques weakening the effectiveness of the prohibition.[404] On the other hand, such a broad interpretation that would lead to an undermining of the national security exception cannot be assumed under EU law either, as this only covers threats to the state as a whole of a certain severity. Recital 24 also supports this limited scope by making clear that systems used outside of these purposes, temporarily or permanently, for other purposes, e.g. for civil or humanitarian purposes, for law enforcement or public security, are covered by the AI Act.[405] The explicit mention of

[402] See Art. 3 (44) AI Act, which defines a 'publicly accessible space' as a 'physical place accessible to an indefinite number of natural persons'; see also Rec. 19: 'Online spaces are not covered, as they are not physical spaces'; see Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 4 Comput. Law Rev. Int. 97, 101.
[403] See Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 63.
[404] Jessie Levano, 'Predictive Policing in the AI Act: meaningful ban or paper tiger?' (*European Law Blog* 5 July 2024) <europeanlawblog.eu> accessed 3 Dec 2024.
[405] See Rec. 24 Sentence 4 AI Act.

law enforcement and public safety already sets a limit to the national security exception.

1.2 Border Control under the AI Act

AI systems used at the EU borders are also to be classified as high-risk systems according to Art. 6 (2) in connection with Annex III No. 7 AI Act (Migration, asylum and border control management). No. 7 provides for 4 use cases for AI systems in this area namely the use of polygraphs or similar instruments; for individual risk assessment; for the examination of asylum applications, visas and residence permits; and finally for the detection, recognition or verification of persons.[406] Although the terms migration, asylum and border control management are not explicitly defined in the AI Act, in accordance with EU law, it refers to all movements of people entering or leaving the EU, either temporarily or permanently, or moving between different EU member states within the EU.[407] The particularly controversial use of emotion detection systems at the borders is not a prohibited practice according to the wording of the final version, as Art. 5 (1) (f) AI Act, which bans these systems, only refers to areas of workplace and education institutions. But this does not result in no regulation at all; instead, they are also considered high-risk systems under Art. 6 (2) AI Act in conjunction with Annex III No. 1 lit. c.

1.3 Access to Education and Social Welfare under the AI Act

AI systems around administration and in particular the systems considered here regarding the allocation of social welfare benefits and access to education can also be

[406] Ludivine Sarah Stewart, 'The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in the shadows?' (2024) 30 Eur. Law J. 122, 128.
[407] See Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 65.

categorized as high-risk systems. AI systems that assess a person's entitlement to social benefits are covered by the area set out in Art. 6 (2) AI Act in conjunction with Annex III No. 5, namely “Access to and enjoyment of essential private services and essential public services and benefits”. No. 5 lit. a of Annex III elaborates on this and refers to systems intended to be used by public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke or reclaim such benefits and services. The question of what is covered by “essential public assistance benefits and services” is specified in Recital 58 by stating in addition to health services, social security benefits, social services for protection in the event of maternity, illness, accidents at work, the need for care or in old age and in the event of loss of employment, as well as social and housing assistance.[408] Therefore, this includes all social benefits that provide protection and support, e.g. in the event of maternity, illness, accidents at work, the need for care or elderly people and job loss, as well as social assistance and housing benefit.[409] If systems such as SyRI are used to detect possible social welfare fraud, this could also be considered a system designed to predict criminal offenses, which, as seen above is prohibited under Art. 5 (1) (d) AI Act. When a risk notification is issued in relation to possible fraud, as with SyRI, it depends on the consequences: if, investigations are carried out in relation to a criminal offense it may be classified as a prohibited practice; but if only a reclaim is issued it is rather a high-risk system because Annex III No. 5 a includes the reduce, revoke, or reclaim of such benefits. Accordingly, AI systems used in social welfare benefits qualify as high-risk systems under the AI Act. In this regard, Recital 58 explicitly refers to the potential

[408] See Rec. 58 Sentence 2 AI Act.
[409] See Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 62.

impact on fundamental rights and the need for protection against possible discrimination when AI systems determine these benefits and services.[410]

Finally, systems in the field of education, and especially in terms of access to education also classify as high-risk systems according to Art. 6 (2) AI Act in conjunction with Annex III No. 3. In this regard, Annex III No. 3 a) explicitly refers to AI systems that are intended to determine access or admission or to assign natural persons to an education system. In addition to a possible violation of the right to education and training, Recital 56 also refers to the risk of discrimination such as the risk of perpetuating historical patterns of discrimination, for example against women, certain age groups, persons with disabilities, or persons of certain racial or ethnic origins or sexual orientation.

1.4 De Minimis threshold for high-risk AI applications

Where an AI system falls under one of the areas of application in Annex III, it is nevertheless not classified as a high-risk system when it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making, Art. 6 (3) AI Act. This “de minimis” clause[411] thus further excludes applications from the area of high-risk systems if they do not pose a great risk to the above-mentioned areas. Art. 6 (3) AI Act states 4 conditions when this is the case, namely when the AI system: is intended to perform a narrowly defined procedural task; is intended to improve the outcome of a previously completed human activity; is intended to recognize decision patterns or deviations from previous decision patterns and is not intended to replace or influence the previously

[410] See Rec. 58 Sentence 3 AI Act.
[411] Ludivine Sarah Stewart, ‘The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in the shadows?’ (2024) 30 Eur. Law J. 122, 128.

completed human assessment without appropriate human review; is intended to perform a preparatory task for an assessment that is relevant for the purposes of the use cases listed in Annex III. Whenever profiling of natural persons is given, however, a system shall always be classified as a high-risk system regardless of this requirement, Art. 6 (3) (d) AI Act. Regarding the use of AI for decision-making, the point that an AI system does not significantly influence the decision-making process is relevant. Recital 53 clarifies that this means an AI system that has no influence on the content and thus on the outcome of the decision-making process, regardless of whether it is a human or an automated decision.[412] This means, although an AI system is used, the decision needs to be based on the will of the person using it.[413] Once again, it is not clear at what point the AI system has no influence on the output and what extent of human control leads to the application of this exception. Additionally questionable is, how decision support systems are to be considered under this. If a provider considers on the basis of its review that the AI system falls under Art. 6 (3) AI Act and is therefore not a high-risk system, the associated obligations do not apply to the provider, who merely has to document the review and register the system in the EU database in accordance with Art. 49 (2) AI Act as stated in Art. 6 (4) AI Act.

1.5 Classification as a responsibility of the operators

Furthermore, the provider of an AI system within the meaning of Art 3 (1) AI Act, is responsible for assessing, as to which of the four risk categories it is subject to.[414] This is because the provider in principle determines the purpose of the AI system as set out in Art. 3 (12) AI Act. Whether an AI system does not fall under the category of high-risk

[412] See Rec. 53 Sentence 2 AI Act.
[413] Paul Voigt and Nils Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024) p. 69.
[414] Francesca Palmiotto, 'When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis' (2024) 25 Ger. Law J. 210, 220.

systems according to Art. 6 (3) AI Act due to the lack of risk of harm to the health, safety or fundamental rights of natural persons is also assessed by the provider itself. Recital 53 merely specifies that a provider who considers that an AI-system does not pose a high risk on the basis of the above conditions should prepare documentation of the assessment before that system is placed on the market or put into service and submit this documentation to the competent national authorities upon request.[415] On the basis of this assessment-based exception, providers can argue their AI system in question does not entail such risks, with the result that the stricter regulations for high-risk AI systems could be circumvented.[416] However, it should also be mentioned that if the market surveillance authority assumes an AI system to be actually high-risk it can assess the system itself and order compliance measures if a high-risk is identified, Art. 80 (1) and (2) AI Act. Besides, the Commission will develop guidelines to specify practical implementation by February 2026, along with a comprehensive list of practical examples of use cases of AI systems that are high-risk and not high-risk, which will hopefully bring clarity to the difficulties of classification, Art. 6 (5) AI Act.

Overall, the systems in consideration here can all be classified as high-risk systems – although, of course, it always depends on the exact design of the system. It is also noticeable that the Recitals often refer to discrimination risks as the reason for the respective classification, suggesting that the problem has been recognized by the legislator. Accordingly, it is now necessary to examine which requirements apply to high-risk systems and to what extent they combat discrimination.

[415] See Rec. 53 AI Act.
[416] Sandra Wachter, 'Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond' (2024) 26 Yale J. L. & Technol. 671, 685; see also Hannah Ruschemeier, 'Art. 6 Einstufungsvorschriften für Hochrisiko-KI-Systeme' in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 90.

2. Requirements for high-risk systems

High-risk AI systems are firstly subject to the mandatory requirements set out in Art. 8-15 AI Act, requirements that are intended to mitigate the risks and ensure a high level of trustworthiness.[417] Further, the regulated actors are subject to obligations under Art. 16-27 AI Act. Most of these requirements are imposed on the provider[418], whereas the deployer has comparatively few obligations.[419] Art. 8 (1) AI Act establishes a general rule of compliance with the requirements while Art. 9 AI Act requires the establishment of a risk management system as a continuous, iterative process that is planned and executed throughout the lifecycle of a high-risk AI system and requires regular systematic reviews and updates. In terms of discrimination risks and their minimization in AI decisions, Art. 10, Art. 14 and Art. 15 of the AI Act are most important in the framework of the requirements for high-risk AI systems.

2.1 Requirements regarding data, Art. 10 AI Act

Art. 10 AI Act sets out requirements with regard to data quality and data governance for training, validation and testing data sets. As seen, data plays a fundamental role in AI and is of particular importance for the quality of an AI system also in decision-making. Recital 67 explicitly refers to the possible risks of discrimination arising from low quality data sets as the reason for these requirements.[420] Art. 10 (1) AI Act firstly stipulates that AI systems shall be developed on the basis of training, validation and test data sets that meet the quality criteria set out in para. 2-5 for ensuring quality. Accordingly, Art. 10 (1) AI Act could be considered a legal obligation within the meaning of Art. 6 (1) (c) GDPR and thus legitimize data processing when personal data is

[417] See Rec. 64 Sentence 1 AI Act.
[418] See e.g. Art. 9-15, Art. 16, Art. 72 AI Act.
[419] See e.g. Art. 26 and Art. 27 AI Act.
[420] Rec. 67 Sentence 1 AI Act.

necessary.[421] But Recital 63 explicitly states that this Regulation is not to be understood as a legal basis for the processing of personal data, unless explicitly provided otherwise in this Regulation, which is why it is questionable whether Art. 10 (1) AI Act really constitutes a general legal basis for data processing.[422] Under Art. 10 (2) AI Act the data used shall be subject to data governance and management practices including the relevant design decisions (lit. a), the procedures for data collection and the origin of the data (lit. b) and relevant data processing procedures (lit. c). With regard to discrimination risks, however, Art. 10 (2) (f), (g) and (3) AI Act are of particular importance. Art. 10 (2) (f) AI Act explicitly provides for an examination about possible biases in high-risk AI systems that may have a negative impact on fundamental rights or lead to discrimination. In accordance with Art. 10 (2) (g) AI Act, appropriate measures to detect, prevent and mitigate these possible biases identified on the basis of the measures under Art. 10 (2) (f) AI Act are also demanded. Art. 10 (3) AI Act additionally requires the training and test data to be relevant, sufficiently representative and as error-free as possible and to have the corresponding statistical properties. Initially, these are highly advantageous provisions for potential discrimination and will help to combat it. But the AI Act does not provide a definition of bias, nor does it specify what are acceptable levels of bias, mitigation strategies or the expected behavior when bias is not detected or mitigated, or even how to prevent bias.[423] Although there are technical ways of identifying bias, these are not standardized and usually do not meet the strict requirements of European non-discrimination law not least because most of these mechanisms were developed in the

[421] For this, e.g. Alexander Golland, 'KI und KI-Verordnung aus datenschutzrechtlicher Sicht' (2024) 18 EuZW 846, 853.
[422] Rec. 63 Sentence 3 AI Act.
[423] Sandra Wachter, 'Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond' (2024) 26 Yale J. L. & Technol. 671, 688.

US.[424] It is therefore a structurally good legal requirement but without specifications on the scope of bias detection, it is not effective. The requirements in Art. 10 (3) AI Act are also rather abstract in nature without any further clarification, so it remains unclear when data sets are representative, error-free and complete.[425] Another limiting factor in the effectiveness of this provision is that these measures must only be appropriate for the intended purpose of the high-risk AI system which leaves room for interpretation.[426] Also important is Art. 10 (5) AI Act which provides an exception for the use of a special category of data for the purpose of ensuring bias detection and correction. This constitutes an explicit exception to the prohibition on processing sensitive data, as it is an EU law provision that provides for specific measures to protect the fundamental rights and interests of data subjects within the meaning of Art. 9 (2) (g) GDPR.[427] In view of the aforementioned use of sensitive data prohibited under the GDPR and the potential that the use of such data can have in the context of bias detection, this is a particularly helpful exception.[428] However, the exception is only given to the extent that it is "strictly necessary", although it is not clear when this is the case but the wording "strictly" rather speaks for a narrow interpretation. Furthermore, this only applies to high-risk systems, meaning that systems that are not considered as high-risk cannot rely on this provision.[429] Also, the exemption solely for sensitive data is not comprehensible; an exemption for 'normal' data too could then have been involved.[430] Although one could argue - a maiore ad minus - that the use of normal

[424] Ibid.
[425] Björn Steinrötter and Jette Markert, 'Datenbezogene Vorgaben der KI-Verordnung' (2024) 9 RDi 400, 402.
[426] Also noting this, see Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 4 Comput. Law Rev. Int. 97, 103.
[427] See Rec. 70 AI Act; see also Alexander Golland, 'KI und KI-Verordnung aus datenschutzrechtlicher Sicht' (2024) 18 EuZW 846, 853.
[428] See VI.3.
[429] Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 4 Comput. Law Rev. Int. 97, 103.
[430] Björn Steinrötter and Jette Markert, 'Datenbezogene Vorgaben der KI-Verordnung' (2024) 9 RDi 400, 404.

data must then be permitted all the more if the use of sensitive data is already permitted. However, this is to be denied due to the wording, which clearly only refers to sensitive data and the assumption of the European legislator that for normal data other processing authorizations apply, e.g. Art. 6 (1), (2) and (3) GDPR.[431]

Overall, Art. 10 AI Act is helpful for combating bias in data sets and potential discriminatory outcomes through AI decision-making, even if only to a limited extend for the reasons mentioned above. However, with further explanations by guidelines and use cases, especially regarding the examination of bias in data sets it is to be expected that Art. 10 AI Act can unfold its potential. It should also be noted that an infringement is subject to penalties under Art. 99 (4) (a) in conjunction with Art. 16 (1) (a) AI Act as this refers to Art. 10 AI Act.

2.2 Requirements for transparency, Art. 13 AI Act

Art. 13 AI Act sets out requirements for the transparency of AI systems. According to this, AI systems shall be designed and developed in such a way that their functioning is sufficiently transparent to enable deployers to interpret the results of the system and use them appropriately. These requirements are intended to address concerns related to opacity and complexity of certain AI systems and to help deployers to comply with their obligations.[432] To achieve this, high-risk AI systems must include instructions for use in a suitable digital format or otherwise that contain concise, complete, accurate and clear information that is relevant, accessible and understandable for the deployers, Art. 13 (2) AI Act. Art. 13 (3) AI Act provides the minimum content for these instructions, such as the identity and contact details of the provider, the characteristics, capabilities

[431] Nadja Braun Binder and Catherine Egli, 'Art. 10 Daten und Daten-Governance' in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 91.
[432] Rec. 72 Sentence 1 AI Act.

and performance limits of the AI system and the human oversight measures referred to in Art. 14 AI Act. According to Art 13 (3) (b) (iii) AI Act, this also includes any known or foreseeable circumstance in connection with the use of the high-risk AI system that may lead to risks to health and safety or fundamental rights, including risks of discrimination.[433] Although it is unlikely that these risks are foreseeable in any way, the obligations of Art. 13 AI Act as a whole can help to ensure that discrimination risks are identified at an early stage because suitable information is available.

Art. 13 AI Act does not initially have any direct impact on the persons concerned but according to Art. 86 (1) AI Act affected persons have the right to receive clear and meaningful explanations from the deployer about the role of the AI system in the decision-making process and the essential elements of the decision taken. There is no concretization as to what the main elements of the decision are, which allows significant freedom of interpretation. Besides, the scope of this right is questionable as it only applies to high-risk systems and under para. 3, it only applies insofar as such a right is not already provided for by other union law.[434] With a view to a possible right to explanation under the GDPR[435], the scope of application could therefore remain very limited. Nevertheless, the right is of great importance for possible discrimination by AI systems, as the ECJ has not yet determined whether there is a right to explanation under the GDPR and the wording of Art. 86 AI Act includes the main elements of the decision taken and thus probably also the most important criteria. Furthermore, as explained above, the corresponding rights to information under the GDPR require a solely automated decision within the meaning of Art. 22 GDPR. No such requirement

[433] Sarah Legner, 'KI-Verordnung und algorithmische Diskriminierung' (2024) 9 RDi 426, 429.
[434] Marieke Luise Merkle, 'Transparenz nach der KI-Verordnung – von der Blackbox zum Open-Book?' (2024) 9 RDi 414, 419.
[435] See VI.4.1.2.

exists under Art. 86 (1) AI Act, which demands a decision “on the basis of the output from a high-risk AI system”, so the criteria are not congruent.[436]

2.3 Requirements on human oversight, Art. 14 AI Act

Art. 14 AI Act requires AI systems to be designed and developed in such a way that they can be effectively monitored by natural persons including suitable human-machine interfaces. According to Art. 14 (2) AI Act, this is aimed at preventing or minimizing risks to health, safety or fundamental rights. Appropriate measures are required which are proportionate to the risks, the level of autonomy and the context of use of the high-risk AI system, Art. 14 (3) AI Act. To enable such human oversight the responsible person must be in a reasonable and proportionate position to understand, e.g. the capacities and limitations (Art. 14 (4) (a)) and especially be aware of the possible tendency to automatically rely or over-rely on the results generated by a high-risk AI system (“automation bias”[437]) when AI systems are used to provide information or recommendations for decisions to be made by natural persons (Art. 14 (4) (b)). Human oversight and the associated ability to intervene is initially helpful for understanding and preventing discrimination. However, Art. 14 AI Act remains vague regarding the exact requirements and partly makes these subject to technical feasibility (Art. 14 (3) (a) AI-Act), making the scope likely to be limited.

2.4 Accuracy, Robustness and Cybersecurity, Art. 15 AI Act

The final requirement particularly relevant for discrimination is Art. 15 AI Act, which sets out the requirements for accuracy, robustness and cybersecurity. An adequate

[436] Sarah Hartmann, 'Art. 86 Recht auf Erläuterung der Entscheidungsfindung im Einzelfall‘ in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 18.
[437] See Nadja Braun Binder and Catherine Egli, ‘Art. 10 Daten und Daten-Governance’ in Mario Martini and Christiane Wendehorst (eds), *KI-VO* (C.H. Beck 2024) para. 51.

level of accuracy, robustness and cybersecurity and consistent performance is required throughout the entire life cycle of the AI system, Art. 15 (1) AI Act. In this respect, the accuracy requirement is likely to include discrimination by AI decisions and thus counteract them.[438] Art. 15 (4) AI Act, which requires the highest possible resilience to errors, faults or inconsistencies also explicitly mentions the risk of feedback loops in learning systems and therefore requires that systems are designed to eliminate or minimize the risk of potentially biased outputs influencing inputs to future operations. As this is a major risk of discrimination, it is beneficial, however, this obligation is only limited to “as far as possible” measures and is lacking in detail. The lack of precision in particular limits their effectiveness against discriminatory AI decision making.[439]

3. Conformity Assessment under the AI Act

A further obligation to be mentioned is the conformity assessment, which is in principle necessary for all high-risk systems to prove that the AI system complies with the requirements of the AI Act. This obligation is imposed on the provider according to Art. 16 (f) in conjunction with Art. 43 AI Act and needs to be done before the AI system is placed on the market or put into operation. For stand-alone high-risk systems under Art. 6 (2) in conjunction with Annex III AI Act (except for biometric high-risk systems), Art. 43 (2) AI Act demands an internal conformity assessment. This is done by a self-assessment of whether the AI system fulfills the relevant essential requirements for high-risk AI systems.[440] According to Art. 43 (2) AI Act, this occurs without the involvement of a notified body. With regard to this internal assessment procedure the

[438] Sarah Legner, 'KI-Verordnung und algorithmische Diskriminierung‘ (2024) 9 RDi 426, 429.
[439] Ibid.
[440] Michael Veale and Frederik Zuiderveen Borgesius, ‘Demystifying the Draft EU Artificial Intelligence Act’ (2021) 4 Comput. Law Rev. Int. 97, 106.

AI Act does not provide any precise specifications or concrete details with the result that it is almost entirely the responsibility of the provider and can be tailored by them.[441] This raises concerns and represents a "major legal loophole" as they themselves confirm compliance with the rules, which they themselves had to interpret due to the many uncertainties.[442] This initially suggests a limited effectiveness of the conformity assessment. Though, the possibility of the Commission under Art. 43 (6) AI Act to modify this in the direction of an external assessment must also be taken into account. Attention also needs to be paid to Recital 125, which justifies the limitation to internal reviews as particularly necessary in the initial phase of the whole AI regulation but also suggests an extension of external control once the work of professional pre-market certifiers has progressed.[443] Accordingly, although there are concerns about the effectiveness of such an internal conformity assessment, at least in the beginning, these are also quite likely to decrease in the future.[444]

4. Fundamental Rights Impact Assessment, Art. 27 AI Act

Finally, probably the most important aspect of the AI Regulation with regard to discrimination by AI systems is the Fundamental Rights Impact Assessment set out in Art. 27 AI Act. This needs to be carried out for stand-alone high-risk AI systems prior to the deployment by bodies governed by public law or private entities that provide public services. The objective of the FRIA is to identify the specific risks to the rights of

[441] Simon Gerdemann, 'Konformitätsbewertung als Kernpflicht der KI-Verordnung' (2024) 31 NJW 2209, 2211.
[442] Sandra Wachter, 'Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond' (2024) 26 Yale J. L. & Technol. 671, 692.
[443] See Rec. 125 Sentence 2, 3 AI Act; see also Simon Gerdemann, 'Konformitätsbewertung als Kernpflicht der KI-Verordnung' (2024) 31 NJW 2209, 2214.
[444] Against the assumption of a third-party assessment in the future on the grounds that Recitals are not legally binding, see Sandra Wachter, 'Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond' (2024) 26 Yale J. L. & Technol. 671, 693.

individuals or groups of individuals likely to be affected and to identify measures to be taken if those risks materialize.[445] According to para. 2, this obligation only applies before the first use and the deployer can rely on already performed FRIAs if a similar use is intended for another AI system. The results of the assessment must be reported to the market surveillance authority, Art. 27 (3) AI Act. Although fundamental rights must be observed in the aforementioned conformity assessment, this is not the main objective of the assessment as is the case with the FRIA.[446] The FRIAs scope is limited in two essential points: firstly, it is only mandatory for the deployers mentioned in para. 1 and secondly, it only applies to certain high-risk systems, which is insufficient with regard to a comprehensive protection of fundamental rights in particular the right to equal treatment.[447] Overall, the FRIA is certainly useful and supportive with regard to possible risks of discrimination but the problems already mentioned elsewhere arise again: Uncertainties about the scope and content of the FRIA and the fact that the deployer itself assesses the extent to which fundamental rights are affected by their AI system.

5. Interim findings – AI Act as a safeguard against discrimination through AI systems

As the first comprehensive AI regulation with direct effect for private and public providers and deployers of AI systems, the AI Act recognizes and addresses the acute problem of discrimination risks. As has been seen, the law itself frequently refers to the protection of fundamental rights and possible risks of discrimination as justification for

[445] See Rec. 96 Sentence 4 AI Act.
[446] Alessandro Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template' (2024) 54 Comput. L. & Secur. Rev. 106020, 5f.
[447] See Alessandro Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template' (2024) 54 Comput. L. & Secur. Rev. 106020, 8.

many provisions in the Recitals. This also becomes clear in some of the requirements and obligations laid down and seems effective and helpful at first glance. On second glance, however, some weak points become apparent.

To begin, the system of categorizing AI systems into the four risk categories is inherently risky or "innovation-friendly".[448] The number of prohibited AI systems in Art. 5 AI Act is quite limited and does include harmful systems but by no means all, meaning that systems with great risks to fundamental rights are only classified as high-risk systems (such as certain predictive policing, emotion recognition and biometric identification AI systems). Part of the problem are missing criteria laid down for determining when a system poses unacceptable risks to society and the individual and is therefore banned.[449] In addition, there is the aforementioned problem of the provider's own classification; due to the additional requirement of Art. 6 (3) AI Act the provider has quite a lot of freedom to demonstrate that its system is not a high-risk AI system, which means almost all requirements and obligations no longer apply. The transparency requirements are comprehensive but as noted not really concerned towards the person. Although the providers must register the AI system in an EU database, accessible to the public as laid down in Art. 49 AI Act this is not publicly available for areas of law enforcement, migration, asylum and border control management, Art. 49 (4) AI Act. This leads to limited transparency towards the public, especially in those areas that are particularly relevant to fundamental rights and non-discrimination.[450] The FRIA is also not ideal when it comes to the protection of

[448] Simon Gerdemann, 'Konformitätsbewertung als Kernpflicht der KI-Verordnung' (2024) 31 NJW 2209, 2211.
[449] See Lilian Edwards, 'Regulating AI in Europe: four problems and four solutions' (*Ada Lovelace Institute* March 2022) p. 11f. <adalovelaceinstitute.org> accessed 3 Dec 2024.
[450] See Ludivine Sarah Stewart, 'The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in the shadows?' (2024) 30 Eur. Law J. 122, 131; see also Ella Jakubowska, Kave Noori, et al., 'EU's AI Act fails to set gold standard for human rights' (*Amnesty International* 3 April 2024) <amnesty.eu> accessed 3 Dec 2024.

fundamental rights because as already mentioned it does not contain any precise requirements and only obliges the provider itself to perform it; moreover, no preventive measures are required, instead it is only necessary to specify what measures will be taken if the risk has actually materialized.[451] The protection of fundamental rights also falls short in terms of enforcement: according to Art. 77 of the AI Act national public authorities or bodies which supervise or enforce the respect of obligations under Union law protecting fundamental rights, including the right to non-discrimination are granted certain powers such as access to any documents created under this regulation. However, this only partially strengthens the powers to enforce the right to non-discrimination[452], nor does it provide any inherent protection for those affected.[453]

Overall, all these problems lead to gaps in protection against discrimination by AI systems. In particular, the many uncertainties due to the vague wording, the limited scope of application of many regulations only to high-risk systems and the lack of involvement of end users and potential affected persons contribute to this. Due to the already mentioned and expected Brussels effect, there is a recognizable risk that insufficient protection of fundamental rights in relation to AI systems will spread and become the general standard.[454]

Despite all these shortcomings, the AI Act is to be welcomed as a whole. The EU's push to establish a framework for AI has turned AI governance and AI policy into a global discourse in the first place, which has certainly had a positive impact.

[451] Ella Jakubowska, Kave Noori, et al., 'EU's AI Act fails to set gold standard for human rights' (*Amnesty International* 3 April 2024) <amnesty.eu> accessed 3 Dec 2024.
[452] On the lack of impact on German public authorities, as this only leads to annex competence without enforcement possibilities, see Sarah Legner, 'KI-Verordnung und algorithmische Diskriminierung' (2024) 9 RDi 426, 431.
[453] Johann Justus Vasel, 'Sieben Sünden und Defizite europäischer KI-Regulierung' (2024) 18 EuZW 829, 831.
[454] See Marco Almada and Anca Radu, 'The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy' (2024) 25 Ger. Law J. 646, 653ff.

Uncertainties regarding interpretation and application will also be minimized as the commission will provide guidelines on various areas as well as the common specifications under Art. 41 AI Act. Another positive aspect is the integration of AI literacy (Art. 4 AI Act) as literacy and awareness are the most important aspects in dealing with AI. It is absolutely necessary to further improve AI literacy, as this is one of the greatest opportunities to master this technique.

VIII. Conclusion

The issue of bias and discrimination in AI-based decision-making is an urgent challenge in the context of advancing technology. As AI systems increasingly influence decisions with significant consequences, the risk builds up that discriminatory outcomes undermine fundamental rights. However, the difficulty of finding universal definitions of discrimination in this context complicates efforts to address these issues effectively. While the GDPR was not explicitly designed to combat AI-related discrimination it does provide some indirect safeguards. The tension between the GDPR and AI highlights the complex interplay between AI advancements and the strict data protection requirements of the GDPR, which can both support and hinder efforts to combat discrimination. Unlike traditional non-discrimination laws, which primarily address discrimination at the outcome level, the GDPR provides upstream protection by regulating the processing of data and emphasizing transparency, accountability and fairness throughout the data lifecycle. While this preventative approach is very valuable in principle, it would be better served by a more active focus of the GDPR on this goal. Though Art. 22 GDPR has a narrow scope of application and thus usually does not apply to decisions made by public authorities in the context of AI, its interpretation and extensions - as found in the SCHUFA ruling - show potential ways to mitigate AI-related

discrimination. The GDPR, despite its supportive elements, is not universally helpful in combating discrimination in AI-powered decision-making. In some cases, its extensive regulatory requirements can unintentionally impede progress and create barriers to innovation and the implementation of effective measures to even combat AI bias. This over-regulation can make it difficult to strike a balance between protecting individual rights and promoting technological advancement for auditing bias. The AI Act partially addresses this problem by introducing exemptions in certain limited areas that allow for more flexibility in the use of personal data. These targeted exemptions aim to mitigate some of the restrictions imposed by the GDPR and allow for more practical solutions without infringing fundamental rights. Besides this, the AI Act represents a significant regulatory effort but is limited in its ability to comprehensively protect individuals from AI bias. Its design as a product safety law prioritizes system-level safety and compliance over the specific protection of individuals, a gap that contrasts with the individual-centric approach of the GDPR. While the current legal framework examined here, namely the GDPR and the AI Act provide some support it is not sufficient to ensure robust protection against discrimination in AI-based decision-making. To close these gaps more targeted legal provisions are needed that emphasize the protection of fundamental rights in different application areas.

What is clearly needed is a mixture of legal approaches, some of which have already been created and technical solutions. Technical measures at the design stage sure help but are not enough, as many of the factors causing discrimination are outside in the 'real world'. This is why auditing while already using the AI system is absolutely necessary.[455] However, the most promising technological development to date appears to be the Explainable AI (XAI) approach, which provides “visibility into how an

[455] See Pauline T. Kim, ‘Auditing Algorithms for Discrimination’ (2017) 166 U. Pa. L. Rev. 189, 196ff.

AI system makes decisions".[456] This transparency allows providers to see how and why an AI has come to a particular conclusion making it easier to identify and correct potential biases. By revealing the decision-making process XAI can help to identify discriminatory patterns or unfair outcomes and ensure that systems do not inadvertently favor or disadvantage certain groups. Thereby, it is at least technically possible for affected individuals to obtain this type of information - assuming they have a right to information.

Nevertheless, AI has significant potential to make decisions without bias or discriminatory outcomes, provided it is developed and trained on robust fairness principles and on the basis of diverse data. Unlike humans, who are inherently influenced by unconscious biases and also act as a 'black box' in their decision-making processes, AI systems can be audited and adjusted to systematically reduce unfairness. Furthermore, AI outperforms human capabilities in terms of speed and scalability, enabling faster and more efficient decision-making processes that can support public administration and government functions. With effective regulation, AI can and will become a transformative tool for equitable and efficient governance balancing its capabilities with the necessary safeguards against potential risks.

[456] Arun Rai, 'Explainable AI: from black box to glass box' (2019) 48 Journal of the Academy of Marketing Science 137, 137f.

Bibliography

‘Artificial intelligence: threats and opportunities’ (*European Parliament* 20 June 2023) <europarl.europa.eu>

Alì GS and R Yu, ‘Artificial Intelligence between Transparency and Secrecy: From the EC Whitepaper to the AIA and Beyond’ (2021) 12 EJLT

Allen R and D Masters, ‘Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making’ (2020) 20 ERA Forum 585

Almada M and A Radu, ‘The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy’ (2024) 25 Ger. Law J. 646

Angwin J and J Larson et al., ‘Machine Bias’ (*propublica* 23 May 2016) <propublica.org>

Antunes HS, PM Freitas, et al. (eds), *Multidisciplinary Perspectives on Artificial Intelligence and the Law* (Springer 2024)

Art 29 Data Protection Working Party, ‘Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679’ 17/EN WP251rev.01

Ballestrem J, U Bär, et al. (eds), *Künstliche Intelligenz: Rechtsgrundlagen und Strategien in der Praxis* (Springer 2020)

Barocas S and AD Selbst, ‘Big Data’s Disparate Impact’ (2016) 104 Calif. Law Rev. 671

Bartoletti I and R Xenidis, *Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination* (Council of Europe 2023)

Bauer F, S Buchberger, et al., 'Machine Learning und die Transparenzanforderungen der DS-GVO‘ (*bitkom* 2018) <bitkom.org>

Boehme-Neßler V, 'Das Ende der Anonymität: Wie Big Data das Datenschutzrecht verändert‘ (2016) 7 Datenschutz und Datensicherheit, 419

Borgesius FZ, ‘Strengthening legal protection against discrimination by algorithms and artificial intelligence’ (2020) 24 Int. J. Hum. Rights 1572

Borgesius FZ, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making* (Council of Europe 2018)

Bradford A, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020)

Chiusi F, S Fischer, et al. (eds), ‘Automating Society’ (*Algorithm Watch, Bertelsmann Stiftung* Oct 2020) <algorithmwatch.org>

Christian G, ‘AI Facial Recognition Technology in the Canadian Immigration System’ (*CILA* 29 August 2023) <cila.co>

Clark J, M Demircan, et al., ‘Europe: The EU AI Act’s relationship with data protection law: key takeaways’ (*DLA Piper* 25 April 2024) <privacymatters.dlapiper.com>

Cormen T, C Leiserson, et al., *Introduction to algorithms* (2nd edn The MIT Press 2022)

Craig P and G de Búrca (eds), *The Evolution of EU Law* (3rd edn Oxford University Press 2021)

Custers B and E Fosch-Villaronga (eds), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice* (T.M.C. Asser Press 2022)

Custers B, T Calders, et al. (eds), *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases* (Springer 2013)

D’Ignazio C and L Klein, *Data Feminism* (MIT Press 2020)

D’Onofrio S, ‘Generative Künstliche Intelligenz – die neue Ära der kreativen Maschinen’ (2024) 61 HMD Praxis der Wirtschaftsinformatik 331

Dacar R, ‘The Essential Facilities Doctrine, Intellectual Property Rights, and Access to Big Data’ (2023) 54 Int. Rev. Intellect. Prop. Compet. Law 1487

Dastin J, ‘Insight - Amazon scraps secret AI recruiting tool that showed bias against women’ (*Reuters* 11 Oct 2018) <reuters.com>

De Gregorio G and P Dunn, ‘The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age’ (2022) 59 Common Mark. Law Rev. 473

Dieterich W, C Mendoza, et al., ‘COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity’ (*Northpointe* 8 July 2016) <go.volarisgroup.com>

Dumbrava C, *Artificial Intelligence at EU borders: Overview of applications and key issues* (European Parliament 2021)

Dürager S, ‘Highlights und Pain Points aus dem "KI-Gesetz" (Teil I)’ (2024) 513 ecolex 898

Edwards L, ‘Regulating AI in Europe: four problems and four solutions’ (*Ada Lovelace Institute* March 2022) <adalovelaceinstitute.org>

Ehmann E and M Selmayr (eds), *Datenschutz-Grundverordnung* (C.H. Beck 2024)

Epping V and C Hillgruber (eds), *BeckOK GG* (58th edn C.H. Beck 2024)

Ertel W, *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung* (5th edn Springer 2021)

European Commission, 'Communication from the Commission to the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Fostering a European approach to Artificial Intelligence' COM (2021) 205

European Commission, 'Communication from the Commission to the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe' COM (2018) 237

European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A Digital Single Market Strategy for Europe' COM (2015) 192 final

European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts' COM (2021) 206

European Commission, 'White Paper on Artificial Intelligence - A European approach to excellence and trust' COM (2020) 65 final

Europol, *AI and Policing: The benefits and challenges of artificial intelligence for law enforcement* (Publications Office of the European Union 2024)

Ferguson A, 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U. Pa. L. Rev. 327

Forti M, 'Addressing Algorithmic Errors in Data-Driven Border Control Procedures' (2024) 25 Ger. Law J. 635

FRA, *#BigData: Discrimination in data-supported decision making* (Publications Office of the European Union 2022)

FRA, *Bias in algorithms - Artificial intelligence and discrimination* (Publications Office of the European Union 2022)

FRA, *Getting the Future Right: Artificial Intelligence and Fundamental Rights* (Publications Office of the European Union 2020)

Friedmann B and H Nissenbaum, 'Bias in Computer Systems' (1996) 14 ACM Trans Inf. Syst. (TOIS) 330

Friis S and J Riley, ‘AI And Machine Learning: Eliminating Algorithmic Bias Is Just the Beginning of Equitable AI’ (2023) Harv. Bus. Rev. <hbr.org>

Fröhlich W and I Spiecker, ‘Können Algorithmen diskriminieren?’ (*Verfblog* 26 Dec 2018) <verfassungsblog.de>

Gerards J and R Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law* (Publications Office of the European Union 2021)

Gerdemann S, ‘Konformitätsbewertung als Kernpflicht der KI-Verordnung’ (2024) 31 NJW 2209

Gola P and D Heckmann (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2022)

Golland A, ‘KI und KI-Verordnung aus datenschutzrechtlicher Sicht’ (2024) 18 EuZW 846

Gupta A, ‘Are Algorithms Sexist?’ (*The New York Times* 15 Nov 2019) <nytimes.com>

Hacker P, ‘Europäische und nationale Regulierung von Künstlicher Intelligenz‘ (2020) 30 NJW 2142

Hacker P, ‘Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law’ (2018) 55 Common Mark. Law Rev. 1143

Häuselmann A and B Custers, ‘Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR’ (2024) 52 Comput. L. & Secur. Rev. 105942

Heaven W, ‘Predictive policing algorithms are racist. They need to be dismantled.’ (2020) MIT Tech. Rev. <technologyreview.com>

Heesen J (ed), *Handbuch Medien-und Informationsethik* (J.B. Metzler 2016)

Hellman D, ‘Measuring Algorithmic Fairness’ (2020) 106 Va. L. Rev. 811

Hoffmeister K, ‘The Dawn of Regulated AI: Analyzing the European AI Act and its Global Impact’ (2024) 2 ZEuS 182

Horstmann J, ‘CJEU: The Rating of a Natural Person’s Creditworthiness by a Credit Rating Agency Constitutes Profiling and Can Be an Automated Decision under Article 22 GDPR’ (2024) 1 Eur. Data Prot. Law Rev. 117

Hüger J, *Künstliche Intelligenz und Diskriminierung* (Nomos 2023)
Huq AZ, ‘A right to a human decision’ (2020) 106 Va. L. Rev. 611

Jackson M, ‘Artificial Intelligence & Algorithmic Bias: The Issues with Technology Reflecting History & Humans’ (2021) 16 J. Bus. & Tech. Law 299

Jakubowska E, K Noori, et al., ‘EU’s AI Act fails to set gold standard for human rights’ (*Amnesty International* 3 April 2024) <amnesty.eu>

Kaminski M, ‘The Right to Explanation, Explained’ (2019) 34 Berkley Tech. L. J. 189

Kaulartz M and T Braegelmann, *Rechtshandbuch Artificial Intelligence and Machine Learning* (C.H. Beck, Vahlen 2020)

Kelleher JD, *Deep Learning* (MIT Press 2019)

Kim P, ‘Auditing Algorithms for Discrimination’ (2017) 166 U. Pa. L. Rev. 189

Kim P, ‘Data-Driven Discrimination at Work’ (2017) 58 Wm. & Mary L. Rev. 857

Knight W, ‘The Dark Secret at the Heart of AI’ (*MIT Technology Review* 11 April 2017) <technologyreview.com>

Kop M, ‘The Right to Process Data for Machine Learning Purposes in the EU’ (2021) 34 Harv. J. L. & Tech. 1

Kreutzer RT, *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmenseigene KI-Journey* (2nd edn Springer 2023)

Kroll J, J Huey, et al., ‘Accountable Algorithms’ (2017) 165 U. Pa. L. Rev. 633

Krönke C, ‘Das europäische KI-Gesetz: Eine Verordnung mit Licht und Schatten’ (2024) 8 NVwZ 529

Kuner C, L Bygrave, et al. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary* (Oxford University Press 2020)

Larson J, S Mattu, et al., ‘How We Analyzed the COMPAS Recidivism Algorithm’ (*propublica* 23 May 2016) <propublica.org>

Legner S, 'KI-Verordnung und algorithmische Diskriminierung‘ (2024) 9 RDi 426

Lerman J, ‘Big Data and its Exclusions’ (2013) 66 Stan. L. Rev. Online 55

Levano J, ‘Predictive Policing in the AI Act: meaningful ban or paper tiger?‘ (*European Law Blog* 5 July 2024) <europeanlawblog.eu>

Lynskey O, ‘Criminal justice profiling and EU data protection law: precarious protection from predictive policing’ (2019) 15 International Journal of Law in Context 162

Malgieri G and G Comande, ‘Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation’ (2017) 7 Int. Data Priv. Law 243

Manning C, ‘Artificial Intelligence Definitions’ *(Stanford HAI Sep 2020)* <hai.stanford.edu>

Mantelero A, ‘The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template’ (2024) 54 Comput. L. & Secur. Rev. 106020

Martini M and C Wendehorst (eds), *KI-VO* (C.H. Beck 2024)

Martini M, 'Algorithmen als Herausforderung für die Rechtsordnung‘ (2017) 21 JZ 1017

Martini M, *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz* (Springer 2019)

Matejek M and S Mäusezahl, ‘Gewöhnliche vs. sensible personenbezogene Daten: Abgrenzung und Verarbeitungsrahmen von Daten gem. Art. 9 DS-GVO’ (2019) 12 ZD 551

Mayson S, ‘Bias in, Bias out’ (2019) 128 Yale L. J. 2218

McCarthy J and P Hayes, ‘Some philosophical problems from the standpoint of artificial intelligence’ (*Stanford Computer Science* 1969) <stanford.edu>

McCarthy J, ML Minsky, et al., ‘A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence’ (1955) <stanford.edu>

Merkle ML, ‘Transparenz nach der KI-Verordnung – von der Blackbox zum Open-Book?’ (2024) 9 RDi 414

Molnar P and L Gill, ‘Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada’s Immigration and Refugee System’ (*IHRP* 2018) <ihrp.law.utoronto.ca>

Mugari I and E Obioha, ‘Predictive Policing and Crime Control in The United States of America and Europe: Trends in a Decade of Research and the Future of Predictive Policing’ (2021) 10 Soc. Sci. <mdpi.com>

Nachbar TB, ‘Algorithmic Fairness, Algorithmic Discrimination’ (2021) 48 Fla. St. U. L. Rev. 509

‘New legislative framework’ (European Commission) <single-market-economy.ec.europe.eu>

Nikolinakos N, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act* (Springer 2023)

Ohm P, ‘Broken Promises of Privacy: responding to the surprising failure of anonymization’ (2010) 57 UCLA L. Rev. 1701

OpenAI ‘Introducing ChatGPT’, (*OpenAI*, 30 Nov 2022) <openai.com>

Orwat C, *Risks of Discrimination through the Use of Algorithms* (Federal Anti-Discrimination Agency 2020)

Paal B and D Pauly (eds), *DS-GVO BDSG* (3rd edn C.H. Beck 2021)

Palmiotto F, ‘When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis’ (2024) 25 Ger. Law J. 210

Park Lee A, ‘Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing’ (2019) UCLA L. Rev. <uclalawreview.org>

Pehlivan CN, ‘The EU Artificial Intelligence (AI) Act: An Introduction’ (2024) 5 Global Privacy Law Review 31

Perez C, *Invisible Women: Exposing data bias in a world designed for men* (Vermilion 2020)

Perry W, B McInnis, et al., *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (Rand 2013)

Radtke T, ‘Das Verhältnis von KI-VO und Art. 22 DS-GVO unter besonderer Berücksichtigung der Schutzzwecke’ (2024) 8 RDi 353

Rai A, ‘Explainable AI: from black box to glass box’ (2019) 48 Journal of the Academy of Marketing Science 137

‘Recommendation of the Council on Artificial Intelligence’ (*OECD* 3 May 2024) <legalinstruments.oecd.org>

Rich E, *Artificial Intelligence* (McGraw-Hill 1983)

Russell JS and P Norvig, *Artificial Intelligence: a modern approach* (4th edn NJ: Pearson 2021)

Sartor G, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (European Parliament 2020)

Scherer MU, ‘Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies’ (2016) 29 Harv. J. L. & Tech. 354

Schneeberger DM, *Machine Learning in der Verwaltung: Rechtsfragen der Black-Box Problematik* (Verlag Österreich 2024)

Selbst AD and J Powles, ‘Meaningful information and the right to explanation’ (2017) 7 Int. Data Priv. Law 233

Selbst AD, ‘The Intuitive Appeal of Explainable Machines’ (2018) 87 Fordham L. Rev. 1085

‘Shaping Europe’s digital future: AI Act’ (*European Commission* 14 Oct 2024) <digital-strategy.ec.europa.eu>

Silfversten E, L Huxtable, et al., ‘Artificial Intelligence – based capabilities for the European Border and Coast Guard’ (FRONTEX 17 May 2021) <frontex.europa.eu>

Spiecker gen. Döhmann I, V Papakonstantinou, et al. (eds), *General Data Protection Regulation* (Nomos 2023)

Steinrötter B and J Markert, ’Datenbezogene Vorgaben der KI-Verordnung’ (2024) 9 RDi 400

Stewart LS, ‘The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in the shadows?’ (2024) 30 Eur. Law J. 122

Surden H, ‘Artificial Intelligence and Law: and Overview’ (2019) 35 Ga. St. U. L. Rev. 1306

Surden H, ‘Machine Learning and Law’ (2014) 89 Wash. L. Rev. 87

Synodinou TE, P Jougleux, et al. (eds), *EU Internet Law: Regulation and Enforcement (*Springer 2017)

Szostek D and M Załucki (eds), *Internet and New Technologies Law: Perspectives and Challenges* (Nomos 2021)

Tene O and J Polonetsky, ‘Privacy in the Age of Big Data’ (2012) 64 Stan. L. Rev. Online 63

Turing A, ‘Computing Machinery and Intelligence‘ (1950) 49 Mind 433

van Bekkum M and FZ Borgesius, ‘Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?’ (2022) 48 Comput. L. & Secur. Rev. 105770

van Dijck G, ‘Predicting Recidivism Risk Meets AI Act’ (2022) 28 Eur. J. Crim. Policy Res. 407

Vasel JJ, ‘Sieben Sünden und Defizite europäischer KI-Regulierung’ (2024) 18 EuZW 829

Vavoula N, ‘Artificial Intelligence (AI) at Schengen Borders: Automated Processing, Algorithmic Profiling and Facial Recognition in the Era of Techno-Solutionism’ (2021) 23 Eur. J. Migr. Law. 457

Veale M and FZ Borgesius, ‘Demystifying the Draft EU Artificial Intelligence Act’ (2021) 4 Comput. Law Rev. Int. 97

Voigt P and N Hullen, *Handbuch KI-Verordnung: FAQ zum EU AI Act* (Springer 2024)

von Welser M, 'Die KI-Verordnung – ein Überblick über das weltweit erste Regelwerk für künstliche Intelligenz' (2024) 15 GRUR-Prax 485

Vrabec H, *Data Subject Rights under the GDPR* (Oxford 2021)

Wachter S, 'Affinity Profiling and Discrimination by Association in Online Behavioral Advertising' (2020) 35 Berkley Tech. L. J. 367

Wachter S, 'Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond' (2024) 26 Yale J. L. & Technol. 671

Wachter S, B Mittelstadt, et al., 'Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR' (2018) 31 Harv. J. L. & Tech. 841

Wachter S, B Mittelstadt, et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 Int. Data Priv. Law 76

Wagener A and C Stark (eds), *Die Digitalisierung des Politischen: Theoretische und praktische Herausforderungen für die Demokratie* (Springer 2023)

Wagner B, P Lopez, et al., 'Der AMS-Algorithmus: Transparenz, Verantwortung und Diskriminierung im Kontext von digitalem staatlichem Handeln' (2020) 2 juridikum 191

Washington A, 'How to argue with an algorithm: Lessons from the COMPAS-ProPublica debate' (2018) 17 Colo. Tech. L. J. 131

Wendehorst C, B Nessler, et al., 'Der Begriff des „KI-Systems" unter der neuen KI-VO' (2024) MMR 605

West SM, 'Discriminating Systems: Gender, Race, and Power in AI' (*AI Now Institute* 1 April 2019) <ainowinstitute.org>

Wimmer B, 'Der AMS-Algorithmus ist ein „Paradebeispiel für Diskriminierung"' (*futurezone* 17 Oct 2018) <futurezone.at>

Wirtz B, J Weyerer, et al., 'Artificial Intelligence and the Public Sector - Applications and Challenges' (2019) 42 Int. J. Public Adm. 596

Wischmeyer T, 'Regulierung intelligenter Systeme' (2018) 143 AöR 1

Wolff H, S Brink, et al. (eds), *BeckOK Datenschutzrecht* (49th edn C.H. Beck 2024)

Wolff J, W Lehr, et al., 'Lessons from GDPR for AI Policymaking' (2024) 27 Va. J. L. & Tech. 1

Yeung K and M Lodge (eds), *Algorithmic Regulation* (Oxford, 2019; online edn)

Zarsky T, 'Transparent Predictions' (2013) 4 Univ. Ill. L. Rev. 1503

Zhou Z and C Firestone, 'Humans can decipher adversarial images' (2019) 10, 1334 Nat. Commun. <nature.com>